OM of: US-08-711-417C-165 to: A_Geneseq_032802:* out_format : pfs

Date: Aug 28, 2002 10:05 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/6228611/runat_28082002_100210_13550/app_query.fasta_1.1639
-DB=A_Geneseq_032802 -QFMT=fastan -SUFFIX=rag -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=6228611_@CGN1_1_164 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-08-711-417C-165
Query length: 1551
Database: A_Geneseq_032802:*
Database sequences: 747574
Database length: 111073796
Search time (sec): 119.550000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation ..
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70971	+	2750.00	4156.83	4.2e-22		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:AAB42333	+	2644.50	3996.44	3.6e-21		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:AAR46964	+	2480.00	3746.08	3.0e-20		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92015	+	2467.00	3728.00	3.6e-20		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72672	+	2467.00	3728.00	3.6e-20		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70964	+	2454.00	3708.24	4.5e-19		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92017	+	2437.00	3681.13	1.3e-19		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72674	+	2437.00	3681.13	1.3e-19		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70966	+	2437.00	3681.13	1.3e-19		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:AAR46965	+	2422.00	3657.33	2.5e-19		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92021	+	2207.50	3333.43	3.3e-17		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72678	+	2207.50	3333.43	3.3e-17		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70970	+	2204.50	3328.87	6.0e-17		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92016	+	1963.00	2962.79	1.6e-15		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72673	+	1963.00	2962.79	1.6e-15		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70965	+	1963.00	2962.79	1.6e-15		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92014	+	1923.50	2902.79	3.5e-15		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:AAR46963	+	1920.50	2898.23	6.4e-15		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72671	+	1920.50	2898.23	6.4e-15		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70963	+	1920.50	2898.23	6.4e-15		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92020	+	1740.00	2626.73	1.1e-13		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72677	+	1740.00	2626.73	1.1e-13		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70969	+	1740.00	2626.73	1.1e-13		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92018	+	1681.00	2535.37	1.1e-13		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72675	+	1681.00	2535.37	1.1e-13		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70967	+	1678.00	2530.81	2.1e-13		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92019	+	1625.00	2450.67	6.2e-12		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72676	+	1625.00	2450.67	6.2e-12		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70968	+	1622.00	2446.11	1.1e-12		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:AAY34146	+	1371.50	2061.75	2.0e-10		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:AAY34148	+	1369.00	2057.95	3.3e-10		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:AAY32714	+	1302.00	1956.54	1.5e-10		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:AAW15574	+	1298.00	1950.46	3.3e-10		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:AAY87745	+	1273.50	1912.68	4.0e-99		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72680	+	1215.00	1832.70	2.6e-94		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:AAY34147	+	1207.50	1813.08	1.5e-93		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72679	+	1010.50	1521.93	5.3e-77		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72681	+	1005.50	1514.24	1.4e-76		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:AAW15575	+	682.50	1024.81	2.9e-49		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:AAY32713	+	682.50	1024.81	2.9e-49		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72692	+	456.00	690.21	3.1e		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72690	+	453.00	685.90	5.4e		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72691	+	448.50	678.44	1.3e		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:AAU16030	+	403.00	604.85	1.1e		
/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:AAU15994	+	349.00	513.23	5.9e		

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70971

seq_documentation_block:

ID	AAW70971 standard; Protein; 516 AA.	
XX		
AC	AAW70971;	
XX		
DT	11-JAN-1999 (first entry)	
XX		
DE	Human Ikaros isoform hIk-1.	
XX		
KW	Ikaros; hIK-1; transcription factor; human; lymphocyte;	
KW	cell differentiation; T cell; cancer; immunodeficiency;	
KW	Alzheimer's disease; therapy; diagnosis.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Region	119..139
FT		/note= "zinc finger motif"
FT	Region	147..167
FT		/note= "zinc finger motif"
FT	Region	175..195
FT		/note= "zinc finger motif"
FT	Region	203..224
FT		/note= "zinc finger motif"
FT	Region	461..481
FT		/note= "zinc finger motif"
FT	Region	489..511
FT		/note= "zinc finger motif"
XX		
PN	CA2194256-A.	
XX		
PD	05-MAR-1998.	
XX		
PF	02-JAN-1997; 97CA-2194256.	
XX		
PR	05-SEP-1996; 96US-0711417.	
XX		
PA	(GEHO) GEN HOSPITAL CORP.	
XX		
PI	Georgopoulos K;	
XX		
DR	WPI; 1998-378292/33.	
DR	N-PSDB; AAV42840.	
XX		
PT	New nucleic acid encoding Ikaros protein involved in early	
PT	differentiation of lymphocytes - existing in several isoforms, and	
PT	related products, used to treat e.g. immune diseases or cancer and	
PT	to control cell differentiation	
XX		
PS	Claim 1; Page 127-129; 158pp; English.	
XX		
CC	This is the amino acid sequence of human Ikaros protein isoform	
CC	hIk-1, deduced from a cDNA clone (see AAV42840) obtained from a	
CC	Jurkat T cell line cDNA library. Native Ikaros is active in the	
CC	early stages of lymphocyte differentiation, binding to and	
CC	activating the CD3-delta gene enhancer (see AAV42804). Proteins	
CC	of the human Ikaros family (see also AAW70964 and AAW70969) are	
CC	isoforms that arise from differential splicing of Ikaros gene	
CC	transcripts, and contain different combinations of zinc fingers.	
CC	They are expressed primarily in T cells in the adult and may play a	
CC	role as a genetic switch regulating entry into the T cell lineage.	
CC	The human and murine sequences (see also AAW70963 and AAW70965-68) are	
CC	very similar. The invention provides Ikaros nucleic acids, vectors	
CC	and host cells expressing Ikaros polypeptides. These can be used	
CC	to treat T and B cell diseases (e.g. immune deficiencies caused by	

```
CC   drugs, radiation or cancer), to control expression of heterologous
CC   genes placed under control of an Ikaros-responsive element, to
CC   treat nervous system diseases (e.g. Alzheimer's disease) and to
CC   modulate cell division, amplification or differentiation, especially
CC   in haematopoietic cells.  Some Ikaros isoforms are antagonistic of
CC   others and may be used to inhibit interaction with DNA sequences.
XX
SQ   Sequence   516 AA;


alignment_scores:
           Quality: 2750.00            Length:    516
             Ratio:  5.329              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-711-417C-165 x AAW70971   ..

 Align seg 1/1  to: AAW70971  from: 1  to: 516

      1 ATGGATGCTGACGAGGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
      1 MetAspAlaAspGluGlyGlnAspMetSerPheSerSerGlyLysGluSe 17

     51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     17 rProProValSerAspThrProAspGluGlyAspGluProMetProIleP 34

    101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAGCTCCAAGAGTGAC 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     34 roGluAspLeuSerThrThrSerGlyGlyGlnGlnSerSerLysSerAsp 50

    151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAA 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ArgValValAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67

    201 TGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAA 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84

    251 TGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGC 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100

    301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT 350
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117

    351 AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    117 uLysCysAspIleCysGlyIleIleCysIleGlyProAsnValLeuMetV 134

    401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnCys 150

    451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCA 500
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    151 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 167

    501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCC 550
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    167 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 184

    551 GGAGGGACGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    184 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 200

    601 CACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    201 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgThrSerLeuGl 217

    651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    217 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG 234

    701 GCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCA 750
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    234 lyThrLeuTyrProValIleLysGluGluThrLysHisSerGluMetAla 250

    751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    251 GluAspLeuCysLysIleGlySerGluArgSerLeuValLeuAspArgLe 267

    801 AGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    267 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 284

    851 GGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAG 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    284 lyAspLysGlyLeuSerAspThrProTyrAspSerAlaThrTyrGluLys 300

    901 GAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    301 GluAsnGluMetMetLysSerHisValMetAspGlnAlaIleAsnAsnAl 317

    951 CATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCC 1000
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    317 aIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProP 334

   1001 CGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC 1050
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    334 roGlyGlySerGluValValProValIleSerProMetTyrGlnLeuHis 350

   1051 AGGCGCTCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAGCGC 1100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    351 ArgArgSerGluGlyThrProArgSerAsnHisSerAlaGlnAspSerAl 367

   1101 CGTGGAGTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGGAGC 1150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    367 aValGluTyrLeuLeuLeuLeuSerLysAlaLysLeuValProSerGluA 384

   1151 GCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAGAGC 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    384 rgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGluSer 400

   1201 AACAACGAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACATCGC 1250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    401 AsnAsnGluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIleAl 417

   1251 CCGACGCGCGCAACGCGTGTCGCTCAAGGAGGAGCACCGCGCCTACGACC 1300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    417 aArgArgAlaGlnArgValSerLeuLysGluGluHisArgAlaTyrAspL 434

   1301 TGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGC 1350
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    434 euLeuArgAlaAlaSerGluAsnSerGlnAspAlaLeuArgValValSer 450

   1351 ACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCGGGTGCT 1400
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    451 ThrSerGlyGluGlnMetLysValTyrLysCysGluHisCysArgValLe 467

   1401 CTTCCTGGATCACGTCATGTACACCATCCACATGGGCTGCCACGGCTTCC 1450
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    467 uPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisGlyPheA 484

   1451 GTGATCCTTTTGAGTGCAACATGTGCGGCTACCACAGCCAGGACCGGTAC 1500
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    484 rgAspProPheGluCysAsnMetCysGlyTyrHisSerGlnAspArgTyr 500

   1501 GAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGCTTCCACATGAGC 1548
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    501 GluPheSerSerHisIleThrArgGlyGluHisArgPheHisMetSer 516

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:AAB42333
```

```
seq_documentation_block:
ID   AAB42333 standard; Protein; 519 AA.
XX
AC   AAB42333;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF2097 polypeptide sequence SEQ ID NO:4194.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   N-PSDB; AAC76542.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 11; Page 3390-3391; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence   519 AA;
```

```
alignment_scores:
            Quality: 2644.50              Length:    519
              Ratio:  5.175                 Gaps:      3
 Percent Similarity:  98.459   Percent Identity: 96.724

alignment_block:
 US-08-711-417C-165 x AAB42333   ..

 Align seg 1/1  to: AAB42333  from: 1  to: 519

      1 ATGGATGCTGACGAGGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG 50
        ||||||||||||||||||||||||||||||      |||||||||||||||
      1 MetAspAlaAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17

     51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
     17 rProProValSerAspThrProAspGluGlyAspGluProMetProIleP 34

    101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAGCTCCAAGAGTGAC 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
     34 roGluAspLeuSerThrThrSerGlyGlyGlnGlnSerSerLysSerAsp 50

    151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAA 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
     51 ArgValValAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67

    201 TGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAA 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
     67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84

    251 TGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGC 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
     84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100

    301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT 350
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117

    351 AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    117 uLysCysAspIleCysGlyIleIleCysIleGlyProAsnValLeuMetV 134

    401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnCys 150

    451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCA 500
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    151 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 167

    501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCC 550
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    167 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 184

    551 GGAGGGACGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    184 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 200

    601 CACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650
        ||||||||||||||||||||||||||||||||||||:::|||||||||
    201 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuGl 217

    651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
    217 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG 234

    701 GCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCA 750
        ||||||||||||||||||||||||||||||||:::|||||||||||||||
    234 lyThrLeuTyrProValIleLysGluGluThrAsnHisSerGluMetAla 250

    751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800
```

```
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 GluAspLeuCysLysIleGlySerGluArgSerLeuValLeuAspArgLe 267

801 AGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
    ||||||||||||||||||||||||||||||||||||||||||||||||||
267 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 284

851 GGGACAAGGGCCTGTCCGACACGCCCTACGAC...AGTGCCACGTACGAG 897
    |||||||||||||||||||||||||||||||   |||||||:::||||||
284 lyAspLysGlyLeuSerAspThrProTyrAspSerSerAlaSerTyrGlu 300

898 AAGGAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAA 947
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 LysGluAsnGluMetMetLysSerHisValMetAspGlnAlaIleAsnAs 317

948 CGCCATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGC 997
    ||||||||||||||||||||||||||||||||||||||||||||||||||
317 nAlaIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrP 334

998 CCCCGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTG 1047
    ||||||||||||||||||||||||||||||||||||||||||||||||||
334 roProGlyGlySerGluValValProValIleSerProMetTyrGlnLeu 350

1048 CACAGG...CGCTCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGA 1094
     |||:::      :::||||||||||||||||||||||||||||||||||
 351 HisLysProLeuAlaGluGlyThrProArgSerAsnHisSerAlaGlnAs 367

1095 CAGCGCCGTGGAGTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCT 1144
     |||||||||||||   ||||||||||||||||||||||||||||||||||
 367 pSerAlaValGluAsnLeuLeuLeuLeuSerLysAlaLysLeuValProS 384

1145 CGGAGCGCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACC 1194
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 384 erGluArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThr 400

1195 GAGAGCAACAACGAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCA 1244
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 401 GluSerAsnAsnGluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHi 417

1245 CATCGCCCGACGCGCGCAACGC...GTGTCGCTCAAGGAGGAGCACCGCG 1291
     |||||||   :::|||::::::   :::||||||||||||||||||||||
 417 sIleAlaProHisAlaArgAsnGlyLeuSerLeuLysGluGluHisArgA 434

1292 CCTACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGC 1341
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 434 laTyrAspLeuLeuArgAlaAlaSerGluAsnSerGlnAspAlaLeuArg 450

1342 GTGGTCAGCACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTG 1391
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 451 ValValSerThrSerGlyGluGlnMetLysValTyrLysCysGluHisCy 467

1392 CCGGGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATGGGCTGCC 1441
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 467 sArgValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysH 484

1442 ACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCACAGCCAG 1491
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 484 isGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHisSerGln 500

1492 GACCGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGCTTCCA 1541
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 501 AspArgTyrGluPheSerSerHisIleThrArgGlyGluHisArgPheHi 517

1542 CATGAGC 1548
     |||||||
 517 sMetSer 519
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:AAR46964

seq_documentation_block:

```
ID   AAR46964 standard; Protein; 537 AA.
XX
AC   AAR46964;
XX
DT   21-OCT-1994  (first entry)
XX
DE   Peptide with Ikaros protein activity.
XX
KW   Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
KW   corpus striatum; regulatory gene.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 536
FT                   /note= "Position is encoded by a stop codon in the
FT                          corresponding nucleotide sequence."
FT
XX
PN   WO9406814-A.
XX
PD   31-MAR-1994.
XX
PF   14-SEP-1993;   93WO-US08743.
XX
PR   14-SEP-1992;   92US-0946233.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1994-118387/14.
DR   N-PSDB; AAQ44980.
XX
PT   T-cell pathway regulatory gene, Ikaros - encodes family of unique
PT   zinc finger proteins, useful for treating immune system disorders
XX
PS   Claim 14; Page 44-46; 112pp; English.
XX
CC   The Ikaros gene encodes a zinc finger protein which can be used in a
CC   therapeutic composition to treat animals with an immune system
CC   disorder.  It may also be used for assessing whether a subject is at
CC   risk for an immune disorder.  It is of particular use in treating a
CC   disorder of the corpus striatum.
XX
SQ   Sequence   537 AA;
```

alignment_scores:

```
            Quality: 2480.00              Length:    487
              Ratio:  5.221                Gaps:      2
 Percent Similarity:  97.536   Percent Identity: 95.688
```

alignment_block:

US-08-711-417C-165 x AAR46964 ..

Align seg 1/1 to: AAR46964 from: 1 to: 537

```
     88 CCCATGCCGATCCCCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAG 137
        |||:::|||:::||||||                              |||:::||
     59 ProLeuProValProGlu.....................GlnArgSe 67

    138 CTCCAAGAGTGACAGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGA 187
        |:::::      |||      ||||||||||||||||||||||||||||
     67 rGlyArgArgProArg......AlaSerAsnValLysValGluThrGlnS 82

    188 GTGATGAAGAGAATGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCG 237
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     82 erAspGluGluAsnGlyArgAlaCysGluMetAsnGlyGluGluCysAla 98

    238 GAGGATTTACGAATGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCA 287
        ||||||||||||||||||||||||||||||||||||||||||||||||||
     99 GluAspLeuArgMetLeuAspAlaSerGlyGluLysMetAsnGlySerHi 115

    288 CAGGGACCAAGGCAGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTC 337
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    115 sArgAspGlnGlySerSerAlaLeuSerGlyValGlyGlyIleArgLeuP 132
```

```
 338 CTAACGGAAAACTAAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCC 387
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 132 roAsnGlyLysLeuLysCysAspIleCysGlyIleIleCysIleGlyPro 148

 388 AATGTGCTCATGGTTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCA 437
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 149 AsnValLeuMetValHisLysArgSerHisThrGlyGluArgProPheGl 165

 438 GTGCAATCAGTGCGGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGC 487
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 165 nCysAsnGlnCysGlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgH 182

 488 ACATCAAGCTGCATTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAAC 537
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 182 isIleLysLeuHisSerGlyGluLysProPheLysCysHisLeuCysAsn 198

 538 TACGCCTGCCGCCGGAGGGACGCCCTCACTGGCCACCTGAGGACGCACTC 587
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 199 TyrAlaCysArgArgArgAspAlaLeuThrGlyHisLeuArgThrHisSe 215

 588 CGTTGGTAAACCTCACAAATGTGGATATTGTGGCCGAAGCTATAAACAGC 637
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 215 rValGlyLysProHisLysCysGlyTyrCysGlyArgSerTyrLysGlnA 232

 638 GAACGTCTTTAGAGGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGC 687
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 232 rgThrSerLeuGluGluHisLysGluArgCysHisAsnTyrLeuGluSer 248

 688 ATGGGCCTTCCGGGCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCA 737
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 249 MetGlyLeuProGlyThrLeuTyrProValIleLysGluGluThrLysHi 265

 738 CAGTGAAATGGCAGAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCG 787
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 265 sSerGluMetAlaGluAspLeuCysLysIleGlySerGluArgSerLeuV 282

 788 TGCTGGACAGACTAGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCT 837
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 282 alLeuAspArgLeuAlaSerAsnValAlaLysArgLysSerSerMetPro 298

 838 CAGAAATTTCTTGGGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGC 887
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 299 GlnLysPheLeuGlyAspLysGlyLeuSerAspThrProTyrAspSerAl 315

 888 CACGTACGAGAAGGAGAACGAAATGATGAAGTCCCACGTGATGGACCAAG 937
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 315 aThrTyrGluLysGluAsnGluMetMetLysSerHisValMetAspGlnA 332

 938 CCATCAACAACGCCATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTG 987
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 332 laIleAsnAsnAlaIleAsnTyrLeuGlyAlaGluSerLeuArgProLeu 348

 988 GTGCAGACGCCCCCGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGAT 1037
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 349 ValGlnThrProProGlyGlySerGluValValProValIleSerProMe 365

1038 GTACCAGCTGCACAGGCGCTCGGAGGGCACCCCGCGCTCCAACCACTCGG 1087
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 365 tTyrGlnLeuHisArgArgSerGluGlyThrProArgSerAsnHisSerA 382

1088 CCCAGGACAGCGCCGTGGAGTACCTGCTGCTGCTCTCCAAGGCCAAGTTG 1137
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 382 laGlnAspSerAlaValGluTyrLeuLeuLeuLeuSerLysAlaLysLeu 398

1138 GTGCCCTCGGAGCGCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCAC 1187
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 399 ValProSerGluArgGluAlaSerProSerAsnSerCysGlnAspSerTh 415

1188 GGACACCGAGAGCAACAACGAGGAGCAGCGCAGCGGTCTTATCTACCTGA 1237
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 415 rAspThrGluSerAsnAsnGluGluGlnArgSerGlyLeuIleTyrLeuT 432

1238 CCAACCACATCGCCCGACGCGCGCAACGCGTGTCGCTCAAGGAGGAGCAC 1287
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 432 hrAsnHisIleAlaArgArgAlaGlnArgValSerLeuLysGluGluHis 448

1288 CGCGCCTACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCT 1337
     |||||||||||||||:::|||||||||||||||||||||||||||||::
 449 ArgAlaTyrAspLeuValArgAlaAlaSerGluAsnSerGlnAspAlaPh 465

1338 CCGCGTGGTCAGCACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAAC 1387
     :|||||||||||||||||||||||||||||||||||||||||||||||||
 465 eArgValValSerThrSerGlyGluGlnMetLysValTyrLysCysGluH 482

1388 ACTGCCGGGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATGGGC 1437
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 482 isCysArgValLeuPheLeuAspHisValMetTyrThrIleHisMetGly 498

1438 TGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCACAG 1487
     |||||||||||||||||||||||||||||:::|||||||||||||||||
 499 CysHisGlyPheArgAspProPheGluCysThrMetCysGlyTyrHisSe 515

1488 CCAGGACCGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGCT 1537
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 515 rGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisArgP 532

1538 TCCACATGAGC 1548
     |||||||:::
 532 heHisMetThr 535
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92015

seq_documentation_block:

```
ID   AAR92015 standard; Protein; 461 AA.
XX
AC   AAR92015;
XX
DT   09-MAY-1996  (first entry)
XX
DE   Human Ikaros protein hIk-1.
XX
KW   Ikaros; transgene; transgenic animal; transgenic mouse; hIk-1;
KW   immunocomprised; immune system disorder; nervous system disorder;
KW   animal model.
XX
OS   Homo sapiens.
XX
PN   WO9604372-A1.
XX
PD   15-FEB-1996.
XX
PF   28-JUL-1995;   95WO-US09345.
XX
PR   29-JUL-1994;   94US-0283300.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1996-129389/13.
DR   N-PSDB; AAT16060.
XX
PT   Transgenic rodent having Ikaros trans-gene (pref. mutated)  - is
PT   severely immuno-compromised and can be used as model to determine
PT   effects of treatment for immune and nervous system disorders
XX
PS   Disclosure; Fig 2; 102pp; English.
XX
CC   An almost full-length cDNA sequence (AAT16060) codes for part
CC   (AAR92015) of the human Ikaros protein, a zinc finger protein that is
CC   a master regulator of haematopoietic differentiation and a major
CC   determinant in lymphocyte specification and development.  Different
CC   isoforms (see AAR92014 and AAR92016-19) of mouse Ikaros have also been
CC   isolated.  Transgenic animals, pref. mice, having a mutated Ikaros
CC   transgene, esp. a mutation that alters the DNA binding domain of the
```

```
CC   Ikaros protein, are used as models to determine the effects of
CC   treatments for immune or nervous system disorders.
XX
SQ   Sequence   461 AA;


alignment_scores:
            Quality: 2467.00              Length:    461
              Ratio:  5.351                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-08-711-417C-165 x AAR92015   ..

 Align seg 1/1  to: AAR92015  from: 1  to: 461

     166 AATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAATGGGCGTGCCTGTGA 215
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       1 AsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAlaCysGl 17

     216 AATGAATGGGGAAGAATGTGCGGAGGATTTACGAATGCTTGATGCCTCGG 265
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      17 uMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspAlaSerG 34

     266 GAGAGAAAATGAATGGCTCCCACAGGGACCAAGGCAGCTCGGCTTTGTCG 315
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      34 lyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAlaLeuSer 50

     316 GGAGTTGGAGGCATTCGACTTCCTAACGGAAAACTAAAGTGTGATATCTG 365
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      51 GlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAspIleCy 67

     366 TGGGATCATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAAAGAAGCC 415
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      67 sGlyIleIleCysIleGlyProAsnValLeuMetValHisLysArgSerH 84

     416 ACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGCGGGGCCTCATTCACC 465
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      84 isThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSerPheThr 100

     466 CAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGGGAGAAGCC 515
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 GlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGluLysPr 117

     516 CTTCAAATGCCACCTCTGCAACTACGCCTGCCGCCGGAGGGACGCCCTCA 565
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     117 oPheLysCysHisLeuCysAsnTyrAlaCysArgArgArgAspAlaLeuT 134

     566 CTGGCCACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATGTGGATAT 615
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     134 hrGlyHisLeuArgThrHisSerValGlyLysProHisLysCysGlyTyr 150

     616 TGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGAGGAACATAAAGAGCG 665
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 CysGlyArgSerTyrLysGlnArgThrSerLeuGluGluHisLysGluAr 167

     666 CTGCCACAACTACTTGGAAAGCATGGGCCTTCCGGGCACACTGTACCCAG 715
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     167 gCysHisAsnTyrLeuGluSerMetGlyLeuProGlyThrLeuTyrProV 184

     716 TCATTAAAGAAGAAACTAAGCACAGTGAAATGGCAGAAGACCTGTGCAAG 765
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     184 alIleLysGluGluThrLysHisSerGluMetAlaGluAspLeuCysLys 200

     766 ATAGGATCAGAGAGATCTCTCGTGCTGGACAGACTAGCAAGTAATGTCGC 815
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 IleGlySerGluArgSerLeuValLeuAspArgLeuAlaSerAsnValAl 217

     816 CAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGACAAGGGCCTGT 865
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     217 aLysArgLysSerSerMetProGlnLysPheLeuGlyAspLysGlyLeuS 234

     866 CCGACACGCCCTACGACAGTGCCACGTACGAGAAGGAGAACGAAATGATG 915
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     234 erAspThrProTyrAspSerAlaThrTyrGluLysGluAsnGluMetMet 250

     916 AAGTCCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTACCTGGG 965
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     251 LysSerHisValMetAspGlnAlaIleAsnAsnAlaIleAsnTyrLeuGl 267

     966 GGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCCCGGGCGGTTCCGAGG 1015
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     267 yAlaGluSerLeuArgProLeuValGlnThrProProGlyGlySerGluV 284

    1016 TGGTCCCGGTCATCAGCCCGATGTACCAGCTGCACAGGCGCTCGGAGGGC 1065
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     284 alValProValIleSerProMetTyrGlnLeuHisArgArgSerGluGly 300

    1066 ACCCCGCGCTCCAACCACTCGGCCCAGGACAGCGCCGTGGAGTACCTGCT 1115
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     301 ThrProArgSerAsnHisSerAlaGlnAspSerAlaValGluTyrLeuLe 317

    1116 GCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGGAGCGCGAGGCGTCCCCGA 1165
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     317 uLeuLeuSerLysAlaLysLeuValProSerGluArgGluAlaSerProS 334

    1166 GCAACAGCTGCCAAGACTCCACGGACACCGAGAGCAACAACGAGGAGCAG 1215
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     334 erAsnSerCysGlnAspSerThrAspThrGluSerAsnAsnGluGluGln 350

    1216 CGCAGCGGTCTTATCTACCTGACCAACCACATCGCCCGACGCGCGCAACG 1265
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     351 ArgSerGlyLeuIleTyrLeuThrAsnHisIleAlaArgArgAlaGlnAr 367

    1266 CGTGTCGCTCAAGGAGGAGCACCGCGCCTACGACCTGCTGCGCGCCGCCT 1315
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     367 gValSerLeuLysGluGluHisArgAlaTyrAspLeuLeuArgAlaAlaS 384

    1316 CCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGCACCAGCGGGGAGCAG 1365
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     384 erGluAsnSerGlnAspAlaLeuArgValValSerThrSerGlyGluGln 400

    1366 ATGAAGGTGTACAAGTGCGAACACTGCCGGGTGCTCTTCCTGGATCACGT 1415
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     401 MetLysValTyrLysCysGluHisCysArgValLeuPheLeuAspHisVa 417

    1416 CATGTACACCATCCACATGGGCTGCCACGGCTTCCGTGATCCTTTTGAGT 1465
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     417 lMetTyrThrIleHisMetGlyCysHisGlyPheArgAspProPheGluC 434

    1466 GCAACATGTGCGGCTACCACAGCCAGGACCGGTACGAGTTCTCGTCGCAC 1515
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     434 ysAsnMetCysGlyTyrHisSerGlnAspArgTyrGluPheSerSerHis 450

    1516 ATAACGCGAGGGGAGCACCGCTTCCACATGAGC 1548
         ||||||||||||||||||||||||||||||||
     451 IleThrArgGlyGluHisArgPheHisMetSer 461

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72672

seq_documentation_block:
ID   AAW72672 standard; Protein; 461 AA.
XX
AC   AAW72672;
XX
DT   14-JAN-1999  (first entry)
XX
DE   Human Ikaros.
XX
KW   CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW   differentiation marker; immune system; corpus striatum; AIDS;
KW   Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   US5824770-A.
```

XX
PD 20-OCT-1998.
XX
PF 05-JUN-1995; 95US-0465590.
XX
PR 02-MAY-1994; 94US-0238212.
PR 14-SEP-1992; 92US-0946233.
PR 14-SEP-1993; 93US-0121438.
PR 05-JUN-1995; 95US-0465590.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Georgopoulos K;
XX
DR WPI; 1998-582621/49.
DR N-PSDB; AAV66969.
XX
PT Ikaros poly:peptide(s) - useful for treating disorders of immune
PT system or corpus striatum
XX
PS Claim 1; Column 55-58; 111pp; English.
XX
CC The present invention describes a purified peptide having at least one
CC of the following properties: (a) it stimulates transcription of a DNA
CC sequence under the control of a delta A element, an NFKB element or an
CC Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
CC a delta A element, an NFKB element or an Ikaros binding oligonucleotide
CC consensus sequence; (c) it competitively inhibits the binding of a
CC naturally occurring Ikaros isoform to any of a delta A element, an NFKB
CC element or an Ikaros binding oligonucleotide consensus sequence; (d) it
CC competitively inhibits Ikaros binding to Ikaros responsive elements; or
CC (e) it inhibits protein-protein interactions of transcriptional complexes
CC formed with naturally occurring Ikaros isoforms. The proteins, provided
CC that they stimulate gene transcription under the control of delta A
CC elements, NFKB elements and/or Ikaros-binding oligonucleotides, bind to
CC delta A elements, NFKB elements and/or Ikaros-binding oligonucleotides,
CC competitively inhibit binding of naturally occurring Ikaros isoforms to
CC delta A elements, NFKB elements and/or Ikaros-binding oligonucleotides,
CC competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
CC inhibit protein-protein interactions of transcriptional complexes with
CC naturally occurring Ikaros isoforms, can be used to treat immune system
CC disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.
CC Alzheimer's disease. The present sequence represents a specifically
CC claimed human Ikaros protein.
XX
SQ Sequence 461 AA;

alignment_scores:
Quality: 2467.00 Length: 461
Ratio: 5.351 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-711-417C-165 x AAW72672 ..

Align seg 1/1 to: AAW72672 from: 1 to: 461

```
166 AATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAATGGGCGTGCCTGTGA 215
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
  1 AsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAlaCysGl 17

216 AATGAATGGGGAAGAATGTGCGGAGGATTTACGAATGCTTGATGCCTCGG 265
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
 17 uMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspAlaSerG 34

266 GAGAGAAAATGAATGGCTCCCACAGGGACCAAGGCAGCTCGGCTTTGTCG 315
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
 34 lyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAlaLeuSer 50

316 GGAGTTGGAGGCATTCGACTTCCTAACGGAAAACTAAAGTGTGATATCTG 365
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
 51 GlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAspIleCy 67

366 TGGGATCATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAAAGAAGCC 415
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
 67 sGlyIleIleCysIleGlyProAsnValLeuMetValHisLysArgSerH 84

416 ACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGCGGGGCCTCATTCACC 465
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
 84 isThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSerPheThr 100

466 CAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGGGAGAAGCC 515
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
101 GlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGluLysPr 117

516 CTTCAAATGCCACCTCTGCAACTACGCCTGCCGCCGGAGGGACGCCCTCA 565
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
117 oPheLysCysHisLeuCysAsnTyrAlaCysArgArgArgAspAlaLeuT 134

566 CTGGCCACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATGTGGATAT 615
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
134 hrGlyHisLeuArgThrHisSerValGlyLysProHisLysCysGlyTyr 150

616 TGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGAGGAACATAAAGAGCG 665
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
151 CysGlyArgSerTyrLysGlnArgThrSerLeuGluGluHisLysGluAr 167

666 CTGCCACAACTACTTGGAAAGCATGGGCCTTCCGGGCACACTGTACCCAG 715
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
167 gCysHisAsnTyrLeuGluSerMetGlyLeuProGlyThrLeuTyrProV 184

716 TCATTAAAGAAGAAACTAAGCACAGTGAAATGGCAGAAGACCTGTGCAAG 765
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
184 alIleLysGluGluThrLysHisSerGluMetAlaGluAspLeuCysLys 200

766 ATAGGATCAGAGAGATCTCTCGTGCTGGACAGACTAGCAAGTAATGTCGC 815
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
201 IleGlySerGluArgSerLeuValLeuAspArgLeuAlaSerAsnValAl 217

816 CAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGACAAGGGCCTGT 865
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
217 aLysArgLysSerSerMetProGlnLysPheLeuGlyAspLysGlyLeuS 234

866 CCGACACGCCCTACGACAGTGCCACGTACGAGAAGGAGAACGAAATGATG 915
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
234 erAspThrProTyrAspSerAlaThrTyrGluLysGluAsnGluMetMet 250

916 AAGTCCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTACCTGGG 965
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
251 LysSerHisValMetAspGlnAlaIleAsnAsnAlaIleAsnTyrLeuGl 267

966 GGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCCCGGGCGGTTCCGAGG 1015
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
267 yAlaGluSerLeuArgProLeuValGlnThrProProGlySerGluV 284

1016 TGGTCCCGGTCATCAGCCCGATGTACCAGCTGCACAGGCGCTCGGAGGGC 1065
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
284 alValProValIleSerProMetTyrGlnLeuHisArgArgSerGluGly 300

1066 ACCCCGCGCTCCAACCACTCGGCCCAGGACAGCGCCGTGGAGTACCTGCT 1115
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
301 ThrProArgSerAsnHisSerAlaGlnAspSerAlaValGluTyrLeuLe 317

1116 GCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGGAGCGCGAGGCGTCCCCGA 1165
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
317 uLeuLeuSerLysAlaLysLeuValProSerGluArgGluAlaSerProS 334

1166 GCAACAGCTGCCAAGACTCCACGGACACCGAGAGCAACAACGAGGAGCAG 1215
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
334 erAsnSerCysGlnAspSerThrAspThrGluSerAsnAsnGluGluGln 350

1216 CGCAGCGGTCTTATCTACCTGACCAACCACATCGCCCGACGCGCGCAACG 1265
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
351 ArgSerGlyLeuIleTyrLeuThrAsnHisIleAlaArgArgAlaGlnAr 367
```

```
1266 CGTGTCGCTCAAGGAGGAGCACCGCGCCTACGACCTGCTGCGCGCCGCCT 1315
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 367 gValSerLeuLysGluGluHisArgAlaTyrAspLeuLeuArgAlaAlaS 384

1316 CCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGCACCAGCGGGGAGCAG 1365
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 384 erGluAsnSerGlnAspAlaLeuArgValValSerThrSerGlyGluGln 400

1366 ATGAAGGTGTACAAGTGCGAACACTGCCGGGTGCTCTTCCTGGATCACGT 1415
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 401 MetLysValTyrLysCysGluHisCysArgValLeuPheLeuAspHisVa 417

1416 CATGTACACCATCCACATGGGCTGCCACGGCTTCCGTGATCCTTTTGAGT 1465
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 417 lMetTyrThrIleHisMetGlyCysHisGlyPheArgAspProPheGluC 434

1466 GCAACATGTGCGGCTACCACAGCCAGGACCGGTACGAGTTCTCGTCGCAC 1515
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 434 ysAsnMetCysGlyTyrHisSerGlnAspArgTyrGluPheSerSerHis 450

1516 ATAACGCGAGGGGAGCACCGCTTCCACATGAGC 1548
     ||||||||||||||||||||||||||||||||
 451 IleThrArgGlyGluHisArgPheHisMetSer 461
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70964

seq_documentation_block:

```
ID   AAW70964 standard; Protein; 461 AA.
XX
AC   AAW70964;
XX
DT   11-JAN-1999  (first entry)
XX
DE   Human Ikaros isoform hIk-1.
XX
KW   Ikaros; hIK-1; transcription factor; human; lymphocyte;
KW   cell differentiation; T cell; cancer; immunodeficiency;
KW   Alzheimer's disease; therapy; diagnosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          64..84
FT                   /note= "zinc finger motif"
FT   Region          92..112
FT                   /note= "zinc finger motif"
FT   Region          120..140
FT                   /note= "zinc finger motif"
FT   Region          148..169
FT                   /note= "zinc finger motif"
FT   Region          406..426
FT                   /note= "zinc finger motif"
FT   Region          434..456
FT                   /note= "zinc finger motif"
XX
PN   CA2194256-A.
XX
PD   05-MAR-1998.
XX
PF   02-JAN-1997;   97CA-2194256.
XX
PR   05-SEP-1996;   96US-0711417.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1998-378292/33.
DR   N-PSDB; AAV42806.
XX
PT   New nucleic acid encoding Ikaros protein involved in early
PT   differentiation of lymphocytes - existing in several isoforms, and
PT   related products, used to treat e.g. immune diseases or cancer and
PT   to control cell differentiation
XX
PS   Claim 7; Page 70-72; 158pp; English.
XX
CC   This is the amino acid sequence of human Ikaros protein isoform
CC   hIk-1, deduced from a cDNA clone (see AAV42806) obtained from a
CC   Jurkat T cell line cDNA library.  Native Ikaros is active in the
CC   early stages of lymphocyte differentiation, binding to and
CC   activating the CD3-delta gene enhancer (see AAV42804).  Proteins
CC   of the human Ikaros family (see also AAW70969 and AAW70971) are
CC   isoforms that arise from differential splicing of Ikaros gene
CC   transcripts, and contain different combinations of zinc fingers.
CC   They are expressed primarily in T cells in the adult and may play a
CC   role as a genetic switch regulating entry into the T cell lineage.
CC   The human and murine sequences (see also AAW70963 and AAW70965-68) are
CC   very similar.  The invention provides Ikaros nucleic acids, vectors
CC   and host cells expressing Ikaros polypeptides.  These can be used
CC   to treat T and B cell diseases (e.g. immune deficiencies caused by
CC   drugs, radiation or cancer), to control expression of heterologous
CC   genes placed under control of an Ikaros-responsive element, to
CC   treat nervous system diseases (e.g. Alzheimer's disease) and to
CC   modulate cell division, amplification or differentiation, especially
CC   in haematopoietic cells.  Some Ikaros isoforms are antagonistic of
CC   others and may be used to inhibit interaction with DNA sequences.
XX
SQ   Sequence   461 AA;
```

alignment_scores:

```
            Quality: 2454.00              Length:    461
              Ratio:  5.335                Gaps:      0
Percent Similarity:  99.783   Percent Identity: 99.349
```

alignment_block:

US-08-711-417C-165 x AAW70964 ..

Align seg 1/1 to: AAW70964 from: 1 to: 461

```
 166 AATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAATGGGCGTGCCTGTGA 215
     ||||||||||||||||||||||||||||||||||||||||||||||||||
   1 AsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAlaCysGl 17

 216 AATGAATGGGGAAGAATGTGCGGAGGATTTACGAATGCTTGATGCCTCGG 265
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  17 uMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspAlaSerG 34

 266 GAGAGAAAATGAATGGCTCCCACAGGGACCAAGGCAGCTCGGCTTTGTCG 315
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  34 lyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAlaLeuSer 50

 316 GGAGTTGGAGGCATTCGACTTCCTAACGGAAAACTAAAGTGTGATATCTG 365
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  51 GlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAspIleCy 67

 366 TGGGATCATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAAAGAAGCC 415
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  67 sGlyIleIleCysIleGlyProAsnValLeuMetValHisLysArgSerH 84

 416 ACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGCGGGGGCCTCATTCACC 465
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  84 isThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSerPheThr 100

 466 CAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGGGAGAAGCC 515
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 101 GlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGluLysPr 117

 516 CTTCAAATGCCACCTCTGCAACTACGCCTGCCGCCGGAGGGACGCCCTCA 565
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 117 oPheLysCysHisLeuCysAsnTyrAlaCysArgArgArgAspAlaLeuT 134

 566 CTGGCCACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATGTGGATAT 615
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 134 hrGlyHisLeuArgThrHisSerValGlyLysProHisLysCysGlyTyr 150
```

```
 616 TGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGAGGAACATAAAGAGCG 665
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 151 CysGlyArgSerTyrLysGlnArgThrSerLeuGluGluHisLysGluAr 167

 666 CTGCCACAACTACTTGGAAAGCATGGGCCTTCCGGGCACACTGTACCCAG 715
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 167 gCysHisAsnTyrLeuGluSerMetGlyLeuProGlyThrLeuTyrProV 184

 716 TCATTAAAGAAGAAACTAAGCACAGTGAAATGGCAGAAGACCTGTGCAAG 765
     ||   ||||||||||||||||||||||||||||||||||||||||||||||
 184 alLysLysGluGluThrLysHisSerGluMetAlaGluAspLeuCysLys 200

 766 ATAGGATCAGAGAGATCTCTCGTGCTGGACAGACTAGCAAGTAATGTCGC 815
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 201 IleGlySerGluArgSerLeuValLeuAspArgLeuAlaSerAsnValAl 217

 816 CAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGACAAGGGCCTGT 865
     ||||||||||||||||||||||||||||||||||||||||||||||||||:
 217 aLysArgLysSerSerMetProGlnLysPheLeuGlyAspLysGlyLeuA 234

 866 CCGACACGCCCTACGACAGTGCCACGTACGAGAAGGAGAACGAAATGATG 915
     ::||||||||||||||||||||||||||||||||||||||||||||||||
 234 laAspThrProTyrAspSerAlaThrTyrGluLysGluAsnGluMetMet 250

 916 AAGTCCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTACCTGGG 965
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 251 LysSerHisValMetAspGlnAlaIleAsnAsnAlaIleAsnTyrLeuGl 267

 966 GGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCCCGGGCGGTTCCGAGG 1015
     |||||||:::||||||||||||||||||||||||||||||||||||||||
 267 yAlaGluAlaLeuArgProLeuValGlnThrProProGlyGlySerGluV 284

1016 TGGTCCCGGTCATCAGCCCGATGTACCAGCTGCACAGGCGCTCGGAGGGC 1065
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 284 alValProValIleSerProMetTyrGlnLeuHisArgArgSerGluGly 300

1066 ACCCCGCGCTCCAACCACTCGGCCCAGGACAGCGCCGTGGAGTACCTGCT 1115
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 301 ThrProArgSerAsnHisSerAlaGlnAspSerAlaValGluTyrLeuLe 317

1116 GCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGGAGCGCGAGGCGTCCCCGA 1165
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 317 uLeuLeuSerLysAlaLysLeuValProSerGluArgGluAlaSerProS 334

1166 GCAACAGCTGCCAAGACTCCACGGACACCGAGAGCAACAACGAGGAGCAG 1215
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 334 erAsnSerCysGlnAspSerThrAspThrGluSerAsnAsnGluGluGln 350

1216 CGCAGCGGTCTTATCTACCTGACCAACCACATCGCCCGACGCGCGCAACG 1265
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 351 ArgSerGlyLeuIleTyrLeuThrAsnHisIleAlaArgArgAlaGlnAr 367

1266 CGTGTCGCTCAAGGAGGAGCACCGCGCCTACGACCTGCTGCGCGCCGCCT 1315
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 367 gValSerLeuLysGluGluHisArgAlaTyrAspLeuLeuArgAlaAlaS 384

1316 CCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGCACCAGCGGGGAGCAG 1365
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 384 erGluAsnSerGlnAspAlaLeuArgValValSerThrSerGlyGluGln 400

1366 ATGAAGGTGTACAAGTGCGAACACTGCCGGGTGCTCTTCCTGGATCACGT 1415
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 401 MetLysValTyrLysCysGluHisCysArgValLeuPheLeuAspHisVa 417

1416 CATGTACACCATCCACATGGGCTGCCACGGCTTCCGTGATCCTTTTGAGT 1465
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 417 lMetTyrThrIleHisMetGlyCysHisGlyPheArgAspProPheGluC 434

1466 GCAACATGTGCGGCTACCACAGCCAGGACCGGTACGAGTTCTCGTCGCAC 1515
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 434 ysAsnMetCysGlyTyrHisSerGlnAspArgTyrGluPheSerSerHis 450

1516 ATAACGCGAGGGGAGCACCGCTTCCACATGAGC 1548
     |||||||||||||||||||||||||||||||||
 451 IleThrArgGlyGluHisArgPheHisMetSer 461
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92017

seq_documentation_block:

```
ID   AAR92017 standard; Protein; 518 AA.
XX
AC   AAR92017;
XX
DT   09-MAY-1996  (first entry)
XX
DE   Murine Ikaros protein mIk-1.
XX
KW   Ikaros; transgene; transgenic animal; transgenic mouse; lymphocyte;
KW   immunocomprised; immune system disorder; nervous system disorder;
KW   animal model; mIk-1.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Domain          119..140
FT.                  /label= F1
FT                   /note= "zinc finger domain F1"
FT   Domain          147..167
FT                   /label= F2
FT                   /note= "zinc finger domain F2"
FT   Domain          175..195
FT                   /label= F3
FT                   /note= "zinc finger domain F3"
FT   Domain          203..224
FT                   /label= F4
FT                   /note= "zinc finger domain F4"
FT   Domain          460..480
FT                   /label= F5
FT                   /note= "zinc finger domain F5"
FT   Domain          491..513
FT                   /label= F6
FT                   /note= "zinc finger domain F6"
XX
PN   WO9604372-A1.
XX
PD   15-FEB-1996.
XX
PF   28-JUL-1995;   95WO-US09345.
XX
PR   29-JUL-1994;   94US-0283300.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1996-129389/13.
DR   N-PSDB; T016062.
XX
PT   Transgenic rodent having Ikaros trans-gene (pref. mutated)  - is
PT   severely immuno-compromised and can be used as model to determine
PT   effects of treatment for immune and nervous system disorders
XX
PS   Disclosure; Fig 4; 102pp; English.
XX
CC   The sequence of 57.5 kDa mouse Ikaros protein mIK-1 (AAR92017) was
CC   deduced from mouse Ikaros cDNA (AAT16062) isolated from a mature
CC   T-cell line E15 library.  Ikaros protein is a master regulator of
CC   hematopoietic differentiation and a major determinant in lymphocyte
CC   differentiation.  Other isoforms of Ikaros (see AAR92014, AAR92016 and
CC   AAR92018-19) arise from differential splicing of Ikaros gene
CC   transcripts.  Transgenic animals, esp. mice, having a mutated Ikaros
CC   transgene, esp. a mutation that alters the DNA binding domain of the
CC   Ikaros protein, are used as models to determine the effects of
CC   treatments for immune or nervous system disorders.
XX
```

SQ Sequence 518 AA;

alignment_scores:
Quality: 2437.00 Length: 521
Ratio: 4.913 Gaps: 6
Percent Similarity: 95.202 Percent Identity: 89.635

alignment_block:
US-08-711-417C-165 x AAR92017 ..

Align seg 1/1 to: AAR92017 from: 1 to: 518

```
   1 ATGGATGCTGACGAGGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG 50
     ||||||:::|||||||||||||||||||||||      ||||||||||||||
   1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17

  51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
     ||||||||||||||||||||||||||||||||||||||||||||||||:::|
  17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34

 101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAGCTCCAAGAGTGAC 150
     |||||||||||||||||||||||||||:::|||||:::||||||||||||
  34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50

 151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAA 200
     |||   :::|||||||||||||||||||||||||||||||||||||||||
  51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67

 201 TGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAA 250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84

 251 TGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGC 300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100

 301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT 350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117

 351 AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400
     ||||||||||||||||||||:::||||||||||||||||||||||||||
 117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134

 401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnSer 150

 451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCA 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 151 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 167

 501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCC 550
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 167 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 184

 551 GGAGGGACGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 184 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 200

 601 CACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650
     ||||||||||||||||||||||||||||||||||||:::|||||||||||
 201 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuGl 217

 651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 217 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG 234

 701 GCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCA 750
     ||   :::   ||||||||||||||||||||:::|||:::||||||||||
 234 ly...ValCysProValIleLysGluGluThrAsnHisAsnGluMetAla 249

 751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800
     ||||||||||||||||||||||:::||||||||||||||||||||||||||
 250 GluAspLeuCysLysIleGlyAlaGluArgSerLeuValLeuAspArgLe 266

 801 AGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 266 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 283

 851 GGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAG 900
     ||||||||   |||||||||   ||||||||||||||:::|||||||||
 283 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 299

 901 GAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
     |||   :::||||||   ||||||||||||||||||||||||||||||||
 300 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl 315

 951 CATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCC 1000
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 315 aIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProP 332

1001 CGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC 1050
     ||||||:::|||||||||||||||||||||||||   ||||||||||||||
 332 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 348

1051 AGG...CGCTCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAG 1097
     :::      |||:::|||   ||||||||||||||||||||||||||||
 349 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp.. 364

1098 CGCCGTGGAGTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGG 1147
      ||||||:::   ||||||||||||||||||||||||   |||   ||||
 365 .AlaValAspAsnLeuLeuLeuLeuSerLysAlaLysSerValSerSerG 381

1148 AGCGCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG 1197
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 381 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 397

1198 AGCAACAACGAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACAT 1247
     ||||||   ||||||||||||||||||||||||||||||||||||||||||
 398 SerAsnAlaGluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIl 414

1248 CGCCCGACGCGCGCAACGC...GTGTCGCTCAAGGAGGAGCACCGCGCCT 1294
     |      :::|||::::::   ::::::||||||||||:::|||||||
 414 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluGlnArgAlaT 431

1295 ACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTG 1344
     ||::::::||||||||||||||||||||||||||||||:::||||||||
 431 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 447

1345 GTCAGCACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCG 1394
     ||||||||||||||||||||:::||||||||||||||||||||||||||
 448 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 464

1395 GGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATG.....G 1435
     |||||||||||||||||||||||||||||||||||||||||||||||||
 464 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 481

1436 GCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCAC 1485
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 481 lyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 497

1486 AGCCAGGACCGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCG 1535
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 498 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 514

1536 CTTCCACATGAGC 1548
     |:::|||:::|||
 514 gTyrHisLeuSer 518
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72674

seq_documentation_block:
ID AAW72674 standard; Protein; 518 AA.
XX

```
AC   AAW72674;
XX
DT   14-JAN-1999  (first entry)
XX
DE   Mouse Ikaros mIk-1.
XX
KW   CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW   differentiation marker; immune system; corpus striatum; AIDS;
KW   Alzheimer's disease.
XX
OS   Mus sp.
XX
PN   US5824770-A.
XX
PD   20-OCT-1998.
XX
PF   05-JUN-1995;   95US-0465590.
XX
PR   02-MAY-1994;   94US-0238212.
PR   14-SEP-1992;   92US-0946233.
PR   14-SEP-1993;   93US-0121438.
PR   05-JUN-1995;   95US-0465590.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1998-582621/49.
DR   N-PSDB; AAV66971.
XX
PT   Ikaros poly:peptide(s) - useful for treating disorders of immune
PT   system or corpus striatum
XX
PS   Claim 1; Column 61-66; 111pp; English.
XX
CC   The present invention describes a purified peptide having at least one
CC   of the following properties: (a) it stimulates transcription of a DNA
CC   sequence under the control of a delta A element, an NFKB element or an
CC   Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
CC   a delta A element, an NFKB element or an Ikaros binding oligonucleotide
CC   consensus sequence; (c) it competitively inhibits the binding of a
CC   naturally occurring Ikaros isoform to any of a delta A element, an NFKB
CC   element or an Ikaros binding oligonucleotide consensus sequence; (d) it
CC   competitively inhibits Ikaros binding to Ikaros responsive elements; or
CC   (e) it inhibits protein-protein interactions of transcriptional complexes
CC   formed with naturally occurring Ikaros isoforms. The proteins, provided
CC   that they stimulate gene transcription under the control of delta A
CC   elements, NFKB elements and/or Ikaros-binding oligonucleotides, bind to
CC   delta A elements, NFKB elements and/or Ikaros-binding oligonucleotides,
CC   competitively inhibit binding of naturally occurring Ikaros isoforms to
CC   delta A elements, NFKB elements and/or Ikaros-binding oligonucleotides,
CC   competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
CC   inhibit protein-protein interactions of transcriptional complexes with
CC   naturally occurring Ikaros isoforms, can be used to treat immune system
CC   disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.
CC   Alzheimer's disease. The present sequence represents a specifically
CC   claimed mouse Ikaros protein.
XX
SQ   Sequence   518 AA;


alignment_scores:
           Quality: 2437.00              Length:    521
             Ratio:  4.913                Gaps:      6
 Percent Similarity:  95.202   Percent Identity: 89.635

alignment_block:
 US-08-711-417C-165 x AAW72674   ..

 Align seg 1/1  to: AAW72674  from: 1  to: 518

       1 ATGGATGCTGACGAGGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG 50
         ||||||:::|||||||||||||||||||||||      |||||||||||||

       1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17

      51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
         ||||||||||||||||||||||||||||||||||||||||||||||:::|
      17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34

     101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAGCTCCAAGAGTGAC 150
         |||||||||||||||||||||||||:::|||||:::|||||||||||||
      34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50

     151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAA 200
         |||   :::|||||||||||||||||||||||||||||||||||||||||
      51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67

     201 TGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAA 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84

     251 TGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100

     301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117

     351 AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400
         |||||||||||||||||||:::|||||||||||||||||||||||||||
     117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134

     401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnSer 150

     451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCA 500
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 167

     501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCC 550
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     167 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 184

     551 GGAGGGACGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     184 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 200

     601 CACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650
         ||||||||||||||||||||||||||||||||||::|||||||||||||
     201 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuGl 217

     651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     217 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG 234

     701 GCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCA 750
         ||   :::   |||||||||||||||||||:::|||:::|||||||||
     234 ly...ValCysProValIleLysGluGluThrAsnHisAsnGluMetAla 249

     751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800
         ||||||||||||||||||:::|||||||||||||||||||||||||||||
     250 GluAspLeuCysLysIleGlyAlaGluArgSerLeuValLeuAspArgLe 266

     801 AGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     266 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 283

     851 GGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAG 900
         ||||||||   |||||||||   ||||||||||||||:::|||||||||
     283 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 299

     901 GAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
         |||   :::||||||   ||||||||||||||||||||||||||||||||
     300 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl 315
```

```
 951 CATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCC 1000
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 315 aIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProP 332

1001 CGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC 1050
     |||||:::|||||||||||||||||||||||||   ||||||||||||||
 332 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 348

1051 AGG...CGCTCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAG 1097
     :::      |||:::|||   ||||||||||||||||||||||||||||
 349 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp.. 364

1098 CGCCGTGGAGTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGG 1147
      ||||||:::   |||||||||||||||||||||||   |||   ||||
 365 .AlaValAspAsnLeuLeuLeuLeuSerLysAlaLysSerValSerSerG 381

1148 AGCGCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG 1197
     |||||||||||||||||||||||||||||||||||||||||||||||||
 381 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 397

1198 AGCAACAACGAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACAT 1247
     ||||||   ||||||||||||||||||||||||||||||||||||||||||
 398 SerAsnAlaGluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIl 414

1248 CGCCCGACGCGCGCAACGC...GTGTCGCTCAAGGAGGAGCACCGCGCCT 1294
     |      :::|||::::::   ::::::||||||||||:::|||||||
 414 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluGlnArgAlaT 431

1295 ACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTG 1344
     ||::::::|||||||||||||||||||||||||||||:::|||||||
 431 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 447

1345 GTCAGCACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCG 1394
     ||||||||||||||||||:::|||||||||||||||||||||||||||
 448 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 464

1395 GGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATG.....G 1435
     |||||||||||||||||||||||||||||||||||||||     |
 464 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 481

1436 GCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCAC 1485
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 481 lyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 497

1486 AGCCAGGACCGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCG 1535
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 498 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 514

1536 CTTCCACATGAGC 1548
     |:::|||:::|||
 514 gTyrHisLeuSer 518
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70966

seq_documentation_block:

```
ID   AAW70966 standard; Protein; 518 AA.
XX
AC   AAW70966;
XX
DT   11-JAN-1999  (first entry)
XX
DE   Mouse Ikaros isoform mIk-1.
XX
KW   Ikaros; mIK-1; transcription factor; mouse; lymphocyte;
KW   cell differentiation; T cell; cancer; immunodeficiency;
KW   Alzheimer's disease; therapy; diagnosis.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          119..139
FT                   /note= "zinc finger motif"
FT   Region          147..167
FT                   /note= "zinc finger motif"
FT   Region          175..195
FT                   /note= "zinc finger motif"
FT   Region          203..224
FT                   /note= "zinc finger motif"
FT   Region          460..480
FT                   /note= "zinc finger motif"
FT   Region          491..513
FT                   /note= "zinc finger motif"
XX
PN   CA2194256-A.
XX
PD   05-MAR-1998.
XX
PF   02-JAN-1997;   97CA-2194256.
XX
PR   05-SEP-1996;   96US-0711417.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1998-378292/33.
DR   N-PSDB; AAV42808.
XX
PT   New nucleic acid encoding Ikaros protein involved in early
PT   differentiation of lymphocytes - existing in several isoforms, and
PT   related products, used to treat e.g. immune diseases or cancer and
PT   to control cell differentiation
XX
PS   Claim 7; Page 75-77; 158pp; English.
XX
CC   This is the amino acid sequence of murine Ikaros protein isoform
CC   mIk-1, deduced from a cDNA clone (see AAV42808) obtained from a
CC   mature murine T cell line E14 library.  Native Ikaros is active
CC   in the early stages of lymphocyte differentiation, binding to and
CC   activating the CD3-delta gene enhancer (see AAV42804).  Proteins
CC   of the murine Ikaros family (see also AAW70963 and AAW70965-68) are
CC   isoforms that arise from differential splicing of Ikaros gene
CC   transcripts, and contain different combinations of zinc fingers.
CC   They are expressed primarily in T cells in the adult and may play a
CC   role as a genetic switch regulating entry into the T cell lineage.
CC   The murine and human sequences (see AAW70964, AAW70969 and AAW70971) are
CC   very similar.  The invention provides Ikaros nucleic acids, vectors
CC   and host cells expressing Ikaros polypeptides.  These can be used
CC   to treat T and B cell diseases (e.g. immune deficiencies caused by
CC   drugs, radiation or cancer), to control expression of heterologous
CC   genes placed under control of an Ikaros-responsive element, to
CC   treat nervous system diseases (e.g. Alzheimer's disease) and to
CC   modulate cell division, amplification or differentiation, especially
CC   in haematopoietic cells.  Some Ikaros isoforms are antagonistic of
CC   others and may be used to inhibit interaction with DNA sequences.
XX
SQ   Sequence   518 AA;
```

alignment_scores:

```
           Quality: 2437.00              Length:    521
             Ratio:  4.913                 Gaps:      6
 Percent Similarity:  95.202   Percent Identity: 89.635
```

alignment_block:

```
 US-08-711-417C-165 x AAW70966   ..

 Align seq 1/1  to: AAW70966  from: 1  to: 518

       1 ATGGATGCTGACGAGGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG 50
         ||||||:::|||||||||||||||||||||      ||||||||||||||
       1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17

      51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
         ||||||||||||||||||||||||||||||||||||||||||||||:::|
```

```
 17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34

101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAGCTCCAAGAGTGAC 150
    ||||||||||||||||||||||||||:::||||||:::|||||||||||
 34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50

151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAA 200
    |||   :::||||||||||||||||||||||||||||||||||||||||
 51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67

201 TGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAA 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84

251 TGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGC 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100

301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117

351 AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400
    |||||||||||||||||||||||:::|||||||||||||||||||||||
117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134

401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
    ||||||||||||||||||||||||||||||||||||||||||||||||||
134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnSer 150

451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCA 500
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 167

501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCC 550
    ||||||||||||||||||||||||||||||||||||||||||||||||||
167 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 184

551 GGAGGGACGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
    ||||||||||||||||||||||||||||||||||||||||||||||||||
184 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 200

601 CACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650
    ||||||||||||||||||||||||||||||||||||:::|||||||||
201 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuGl 217

651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700
    ||||||||||||||||||||||||||||||||||||||||||||||||||
217 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG 234

701 GCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCA 750
    ||   :::   |||||||||||||||||||:::|||:::|||||||||
234 ly...ValCysProValIleLysGluGluThrAsnHisAsnGluMetAla 249

751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800
    ||||||||||||||||||||:::|||||||||||||||||||||||||||
250 GluAspLeuCysLysIleGlyAlaGluArgSerLeuValLeuAspArgLe 266

801 AGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
    ||||||||||||||||||||||||||||||||||||||||||||||||||
266 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 283

851 GGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAG 900
    ||||||||   |||||||||   ||||||||||||||:::|||||||||
283 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 299

901 GAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
    |||   :::||||||   ||||||||||||||||||||||||||||||||
300 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl 315

951 CATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCC 1000
    ||||||||||||||||||||||||||||||||||||||||||||||||||
315 aIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProP 332

1001 CGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC 1050
     |||||:::|||||||||||||||||||||||||   ||||||||||||||
 332 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 348

1051 AGG...CGCTCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAG 1097
     :::      |||:::|||   ||||||||||||||||||||||||||||
 349 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp.. 364

1098 CGCCGTGGAGTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGG 1147
      ||||||:::   |||||||||||||||||||||||   |||   ||||
 365 .AlaValAspAsnLeuLeuLeuLeuSerLysAlaLysSerValSerSerG 381

1148 AGCGCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG 1197
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 381 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 397

1198 AGCAACAACGAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACAT 1247
     ||||||   ||||||||||||||||||||||||||||||||||||||||
 398 SerAsnAlaGluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIl 414

1248 CGCCCGACGCGCGCAACGC...GTGTCGCTCAAGGAGGAGCACCGCGCCT 1294
     |      :::|||::::::   ::::::||||||||||||:::|||||||
 414 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluGlnArgAlaT 431

1295 ACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTG 1344
     ||::::::|||||||||||||||||||||||||||||||||:::||||||
 431 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 447

1345 GTCAGCACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCG 1394
     ||||||||||||||||||||:::|||||||||||||||||||||||||||
 448 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 464

1395 GGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATG.....G 1435
     ||||||||||||||||||||||||||||||||||||||||||   |
 464 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 481

1436 GCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCAC 1485
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 481 lyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 497

1486 AGCCAGGACCGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCG 1535
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 498 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 514

1536 CTTCCACATGAGC 1548
     |:::|||:::|||
 514 gTyrHisLeuSer 518
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:AAR46965

seq_documentation_block:

```
ID   AAR46965 standard; Protein; 568 AA.
XX
AC   AAR46965;
XX
DT   21-OCT-1994  (first entry)
XX
DE   Ikaros zinc finger protein isoform IK-1.
XX
KW   Ikaros; zinc finger; protein; immune disorder; therapy; treatment;
KW   corpus striatum; regulatory gene.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1..53
FT                   /label= Exons 1/2.
FT   Region          54..141
FT                   /label= Exon 3.
FT   Region          142..247
FT                   /label= Exon 4.
FT   Region          248..288
```

```
FT                   /label= Exon 5.
FT   Region          289..333
FT                   /label= Exon 6.
FT   Region          334..568
FT                   /label= Exon 7.
XX
PN   WO9406814-A.
XX
PD   31-MAR-1994.
XX
PF   14-SEP-1993;   93WO-US08743.
XX
PR   14-SEP-1992;   92US-0946233.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1994-118387/14.
XX
PT   T-cell pathway regulatory gene, Ikaros - encodes family of unique
PT   zinc finger proteins, useful for treating immune system disorders
XX
PS   Claim 14; Figure 4; 112pp; English.
XX
CC   The Ikaros gene encodes a zinc finger protein which can be used in a
CC   therapeutic composition to treat animals with an immune system
CC   disorder.  It may also be used for assessing whether a subject is at
CC   risk for an immune disorder.  It is of particular use in treating a
CC   disorder of the corpus striatum.
XX
SQ   Sequence   568 AA;
```

```
alignment_scores:
            Quality: 2422.00              Length:    571
              Ratio:  4.863                Gaps:      7
 Percent Similarity:  87.215   Percent Identity: 82.137

alignment_block:
 US-08-711-417C-165 x AAR46965   ..

 Align seg 1/1  to: AAR46965  from: 1  to: 568

       1 ATGGATGCTGACGAGGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG 50
         ||||||:::|||||||||||||||||||||      ||||||||||||||
       1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17

      51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
         ||||||||||||||||||||||||||||||||||||||||||||||:::|
      17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34

     101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAGCTCCAAGAGTGAC 150
         |||||||||||||||||||||||:::||||||:::||||||||||||||
      34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50

     151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAA 200
         |||   :::|||||||||||||||||||||||||||||||||||||||||
      51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67

     201 TGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAA 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84

     251 TGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100

     300 ..................................................  300

     101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117

     300 ..................................................  300

     117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134

     300 ..................................................  300

     134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnCys 150

     301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 167

     351 AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400
         ||||||||||||||||||||||:::|||||||||||||||||||||||||
     167 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 184

     401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     184 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnCys 200

     451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCCGGCACATCAAGCTGCA 500
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     201 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 217

     501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCC 550
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     217 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 234

     551 GGAGGGACGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     234 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSerValGlyLysPro 250

     601 CACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650
         ||||||||||||||||||||||||||||||||||||||:::|||||||||
     251 HisLysCysGlyTyrCysGlyArgSerTyrLysGlnArgSerSerLeuGl 267

     651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     267 uGluHisLysGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProG 284

     701 GCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCA 750
         ||   :::|||||||||||||||||||||||||:::|||:::|||||||||
     284 ly...MetTyrProValIleLysGluGluThrAsnHisAsnGluMetAla 299

     751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800
         ||||||||||||||||||||:::|||||||||||||||||||||||||||
     300 GluAspLeuCysLysIleGlyAlaGluArgSerLeuValLeuAspArgLe 316

     801 AGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     316 uAlaSerAsnValAlaLysArgLysSerSerMetProGlnLysPheLeuG 333

     851 GGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAG 900
         ||||||||   |||||||||   ||||||||||||||:::|||||||||
     333 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 349

     901 GAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
         |||   :::||||||   ||||||||||||||||||||||||||||||||
     350 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl 365

     951 CATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCC 1000
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     365 aIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProP 382

    1001 CGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC 1050
         |||||:::|||||||||||||||||||||||||   ||||||||||||||
     382 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 398

    1051 AGG...CGCTCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAG 1097
         :::      |||:::|||   |||||||||||||||||||||||||||||
     399 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp.. 414

    1098 CGCCGTGGAGTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGG 1147
         |||||||:::   ||||||||||||||||||||||||   |||   ||||
```

```
 415 .AlaValAspAsnLeuLeuLeuLeuSerLysAlaLysSerValSerSerG 431

1148 AGCGCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG 1197
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 431 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 447

1198 AGCAACAACGAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACAT 1247
     ||||||   ||||||||||||||||||||||||||||||||||||||||
 448 SerAsnAlaGluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIl 464

1248 CGCCCGACGCGCGCAACGC...GTGTCGCTCAAGGAGGAGCACCGCGCCT 1294
     |      :::|||:::::::   :::::::|||||||||||||:::|||||||
 464 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluGlnArgAlaT 481

1295 ACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTG 1344
     ||::::::|||||||||||||||||||||||||||||||:::|||||||
 481 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 497

1345 GTCAGCACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCG 1394
     ||||||||||||||||||:::|||||||||||||||||||||||||||||
 498 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 514

1395 GGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATG.........G 1435
     ||||||||||||||||||||||||||||||||||||||||         |
 514 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 531

1436 GCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCAC 1485
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 531 lyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 547

1486 AGCCAGGACCGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCG 1535
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 548 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 564

1536 CTTCCACATGAGC 1548
     |:::|||:::|||
 564 gTyrHisLeuSer 568
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92021

seq_documentation_block:

```
ID   AAR92021 standard; Protein; 470 AA.
XX
AC   AAR92021;
XX
DT   09-MAY-1996  (first entry)
XX
DE   Ikaros protein.
XX
KW   Ikaros; transgene; transgenic animal; transgenic mouse; lymphocyte;
KW   immunocomprised; immune system disorder; nervous system disorder;
KW   animal model.
XX
OS   Not specified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..2
FT                   /note= "unidentified amino acids"
FT   Misc-difference 74
FT                   /note= "unidentified amino acid"
FT   Misc-difference 163
FT                   /note= "unidentified amino acid"
FT   Misc-difference 184..186
FT                   /note= "unidentified amino acids"
FT   Misc-difference 194
FT                   /note= "unidentified amino acid"
FT   Misc-difference 196
FT                   /note= "unidentified amino acid"
FT   Misc-difference 207
FT                   /note= "unidentified amino acid"
FT   Misc-difference 236
FT                   /note= "unidentified amino acid"
FT   Misc-difference 240
FT                   /note= "unidentified amino acid"
FT   Misc-difference 246
FT                   /note= "unidentified amino acid"
FT   Misc-difference 251..252
FT                   /note= "unidentified amino acids"
FT   Misc-difference 255
FT                   /note= "unidentified amino acid"
FT   Misc-difference 261
FT                   /note= "unidentified amino acid"
FT   Misc-difference 285
FT                   /note= "unidentified amino acid"
FT   Misc-difference 300..302
FT                   /note= "unidentified amino acids"
FT   Misc-difference 304
FT                   /note= "unidentified amino acid"
FT   Misc-difference 306
FT                   /note= "unidentified amino acid"
FT   Misc-difference 316
FT                   /note= "unidentified amino acid"
FT   Misc-difference 319..320
FT                   /note= "unidentified amino acids"
FT   Misc-difference 329
FT                   /note= "unidentified amino acid"
FT   Misc-difference 331
FT                   /note= "unidentified amino acid"
FT   Misc-difference 352
FT                   /note= "unidentified amino acid"
FT   Misc-difference 367..369
FT                   /note= "unidentified amino acids"
FT   Misc-difference 371..375
FT                   /note= "unidentified amino acids"
FT   Misc-difference 380
FT                   /note= "unidentified amino acid"
FT   Misc-difference 384..385
FT                   /note= "undientified amino acids"
FT   Misc-difference 397
FT                   /note= "unidentified amino acid"
FT   Misc-difference 407
FT                   /note= "unidentified amino acid"
FT   Misc-difference 430..432
FT                   /note= "unidentified amino acids"
FT   Misc-difference 467
FT                   /note= "unidentified amino acid"
FT   Misc-difference 469
FT                   /note= "unidentified amino acid"
XX
PN   WO9604372-A1.
XX
PD   15-FEB-1996.
XX
PF   28-JUL-1995;   95WO-US09345.
XX
PR   29-JUL-1994;   94US-0283300.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1996-129389/13.
XX
PT   Transgenic rodent having Ikaros trans-gene (pref. mutated)  - is
PT   severely immuno-compromised and can be used as model to determine
PT   effects of treatment for immune and nervous system disorders
XX
PS   Disclosure; Page 75-76; 102pp; English.
XX
CC   The sequence of an Ikaros protein (AAR92021) is provided in the
CC   specification.  Ikaros protein is a master regulator of
CC   hematopoietic differentiation and a major determinant in lymphocyte
CC   differentiation.  Isoforms of Ikaros (see AAR92014-19) arise from
CC   differential splicing of Ikaros gene transcripts.  Transgenic animals,
CC   pref. mice, having a mutated Ikaros transgene, esp. a mutation that
CC   alters the DNA binding domain of the Ikaros protein, are used as
```

```
CC   models to determine the effects of treatments for immune or nervous
CC   system disorders.
XX
SQ   Sequence   470 AA;


alignment_scores:
           Quality: 2207.50              Length:    468
             Ratio:  5.098                Gaps:      3
 Percent Similarity:  92.521   Percent Identity: 90.385

alignment_block:
 US-08-711-417C-165 x AAR92021   ..

 Align seg 1/1  to: AAR92021  from: 1  to: 470

     160 GCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAATGGGCGTGC 209
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       3 AlaSerAsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAl 19

     210 CTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAATGCTTGATG 259
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      19 aCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspA 36

     260 CCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGCAGCTCGGCT 309
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      36 laSerGlyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAla 52

     310 TTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACTAAAGTGTGA 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      53 LeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAs 69

     360 TATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAAA 409
         ||||||||||||   ||||||||||||||||||||||||||||||||||
      69 pIleCysGlyIle***CysIleGlyProAsnValLeuMetValHisLysA 86

     410 GAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGCGGGGCCTCA 459
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      86 rgSerHisThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSer 102

     460 TTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGGGA 509
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 PheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGl 119

     510 GAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCCGGAGGGACG 559
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     119 uLysProPheLysCysHisLeuCysAsnTyrAlaCysArgArgArgAspA 136

     560 CCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATGT 609
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     136 laLeuThrGlyHisLeuArgThrHisSerValGlyLysProHisLysCys 152

     610 GGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGAGGAACATAA 659
         ||||||||||||||||||||||||||||||:::||||||||||||||||||
     153 GlyTyrCysGlyArgSerTyrLysGlnArg***SerLeuGluGluHisLy 169

     660 AGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGGGCACACTGT 709
         ||||||||||||||||||||||||||||||||||||||||||||||:::
     169 sGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProGly******* 186

     710 ACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCAGAAGACCTG 759
           ||||||||||||||||||||   |||:::||||||||||||||||||||
     186 **ProValIleLysGluGluThr***His***GluMetAlaGluAspLeu 202

     760 TGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACTAGCAAGTAA 809
         ||||||||||||:::||||||||||||||||||||||||||||||||||
     203 CysLysIleGly***GluArgSerLeuValLeuAspArgLeuAlaSerAs 219

     810 TGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGACAAGG 859
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     219 nValAlaLysArgLysSerSerMetProGlnLysPheLeuGlyAspLys* 236

     860 GCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAGGAGAACGAA 909
```

```
         |||||||||:::|||||||||||||||:::||||||||||||
     236 **LeuSerAsp***ProTyrAspSerAla***TyrGluLysGlu****** 252

     910 ATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTA 959
         ||||||   ||||||||||||||   |||||||||||||||||||||||||
     253 MetMet***SerHisValMetAsp***AlaIleAsnAsnAlaIleAsnTy 269

     960 CCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCCCGGGCGGTT 1009
         ||||||||||||||||||||||||||||||||||||||||||||||||   |
     269 rLeuGlyAlaGluSerLeuArgProLeuValGlnThrProProGly***S 286

    1010 CCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC...AGGCGC 1056
         ||||||||||||||||||||||||||||||||||||||||
     286 erGluValValProValIleSerProMetTyrGlnLeuHis********* 302

    1057 TCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAGCGCCGTGGA 1106
         |||   |||:::|||||||||||||||||||||||||||:::||||||
     303 Ser***Gly***ProArgSerAsnHisSerAlaGlnAsp***AlaVal** 319

    1107 GTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGGAGCGCGAGG 1156
              ||||||||||||||||||||||   |||   ||||||||||||||
     319 ****LeuLeuLeuLeuSerLysAlaLys***Val***SerGluArgGluA 336

    1157 CGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAGAGCAACAAC 1206
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     336 laSerProSerAsnSerCysGlnAspSerThrAspThrGluSerAsn*** 352

    1207 GAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACATCGCCCGACG 1256
         ||||||||||||||||||||||||||||||||||||||||||||||:::
     353 GluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIle******** 369

    1257 CGCG...CAACGCGTGTCGCTCAAGGAGGAGCACCGCGCCTACGACCTGC 1303
          |||            :::||||||||||||   |||||||||     |
     369 *Ala***************LeuLysGluGlu***ArgAlaTyr******L 386

    1304 TGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGCACC 1353
         ||||||||||||||||||||||||||||||   ||||||||||||||||
     386 euArgAlaAlaSerGluAsnSerGlnAspAla***ArgValValSerThr 402

    1354 AGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCGGGTGCTCTT 1403
         ||||||||||||   ||||||||||||||||||||||||||||||||||
     403 SerGlyGluGln***LysValTyrLysCysGluHisCysArgValLeuPh 419

    1404 CCTGGATCACGTCATGTACACCATCCACATG.........GGCTGCCACG 1444
         ||||||||||||||||||||||||||||||         |||||||||
     419 eLeuAspHisValMetTyrThrIleHisMet*********GlyCysHisG 436

    1445 GCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCACAGCCAGGAC 1494
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     436 lyPheArgAspProPheGluCysAsnMetCysGlyTyrHisSerGlnAsp 452

    1495 CGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGCTTCCACAT 1544
         ||||||||||||||||||||||||||||||||||||||||||   |||
     453 ArgTyrGluPheSerSerHisIleThrArgGlyGluHisArg***His** 469

    1545 GAGC 1548
          |||
     469 *Ser 470
```

```
seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72678

seq_documentation_block:
ID   AAW72678 standard; Protein; 470 AA.
XX
AC   AAW72678;
XX
DT   14-JAN-1999  (first entry)
XX
DE   Ikaros protein general formula.
XX
KW   CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW   differentiation marker; immune system; corpus striatum; AIDS;
KW   Alzheimer's disease.
```

```
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "any amino acid"
FT   Misc-difference 2
FT                   /note= "any amino acid"
FT   Misc-difference 74
FT                   /note= "any amino acid"
FT   Misc-difference 163
FT                   /note= "any amino acid"
FT   Misc-difference 184
FT                   /note= "any amino acid"
FT   Misc-difference 185
FT                   /note= "any amino acid"
FT   Misc-difference 186
FT                   /note= "any amino acid"
FT   Misc-difference 194
FT                   /note= "any amino acid"
FT   Misc-difference 196
FT                   /note= "any amino acid"
FT   Misc-difference 207
FT                   /note= "any amino acid"
FT   Misc-difference 236
FT                   /note= "any amino acid"
FT   Misc-difference 240
FT                   /note= "any amino acid"
FT   Misc-difference 246
FT                   /note= "any amino acid"
FT   Misc-difference 251
FT                   /note= "any amino acid"
FT   Misc-difference 252
FT                   /note= "any amino acid"
FT   Misc-difference 255
FT                   /note= "any amino acid"
FT   Misc-difference 261
FT                   /note= "any amino acid"
FT   Misc-difference 285
FT                   /note= "any amino acid"
FT   Misc-difference 300
FT                   /note= "any amino acid"
FT   Misc-difference 301
FT                   /note= "any amino acid"
FT   Misc-difference 302
FT                   /note= "any amino acid"
FT   Misc-difference 304
FT                   /note= "any amino acid"
FT   Misc-difference 306
FT                   /note= "any amino acid"
FT   Misc-difference 316
FT                   /note= "any amino acid"
FT   Misc-difference 319
FT                   /note= "any amino acid"
FT   Misc-difference 320
FT                   /note= "any amino acid"
FT   Misc-difference 329
FT                   /note= "any amino acid"
FT   Misc-difference 331
FT                   /note= "any amino acid"
FT   Misc-difference 352
FT                   /note= "any amino acid"
FT   Misc-difference 367
FT                   /note= "any amino acid"
FT   Misc-difference 368
FT                   /note= "any amino acid"
FT   Misc-difference 369
FT                   /note= "any amino acid"
FT   Misc-difference 371
FT                   /note= "any amino acid"
FT   Misc-difference 372
FT                   /note= "any amino acid"
FT   Misc-difference 373
FT                   /note= "any amino acid"
FT   Misc-difference 374
FT                   /note= "any amino acid"
FT   Misc-difference 375
FT                   /note= "any amino acid"
FT   Misc-difference 380
FT                   /note= "any amino acid"
FT   Misc-difference 384
FT                   /note= "any amino acid"
FT   Misc-difference 385
FT                   /note= "any amino acid"
FT   Misc-difference 397
FT                   /note= "any amino acid"
FT   Misc-difference 407
FT                   /note= "any amino acid"
FT   Misc-difference 430
FT                   /note= "any amino acid"
FT   Misc-difference 431
FT                   /note= "any amino acid"
FT   Misc-difference 432
FT                   /note= "any amino acid"
FT   Misc-difference 467
FT                   /note= "any amino acid"
FT   Misc-difference 469
FT                   /note= "any amino acid"
XX
PN   US5824770-A.
XX
PD   20-OCT-1998.
XX
PF   05-JUN-1995;   95US-0465590.
XX
PR   02-MAY-1994;   94US-0238212.
PR   14-SEP-1992;   92US-0946233.
PR   14-SEP-1993;   93US-0121438.
PR   05-JUN-1995;   95US-0465590.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1998-582621/49.
XX
PT   Ikaros poly:peptide(s) - useful for treating disorders of immune
PT   system or corpus striatum
XX
PS   Claim 1; Column 127-130; 111pp; English.
XX
CC   The present invention describes a purified peptide having at least one
CC   of the following properties: (a) it stimulates transcription of a DNA
CC   sequence under the control of a delta A element, an NFKB element or an
CC   Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
CC   a delta A element, an NFKB element or an Ikaros binding oligonucleotide
CC   consensus sequence; (c) it competitively inhibits the binding of a
CC   naturally occurring Ikaros isoform to any of a delta A element, an NFKB
CC   element or an Ikaros binding oligonucleotide consensus sequence; (d) it
CC   competitively inhibits Ikaros binding to Ikaros responsive elements; or
CC   (e) it inhibits protein-protein interactions of transcriptional complexes
CC   formed with naturally occurring Ikaros isoforms. The proteins, provided
CC   that they stimulate gene transcription under the control of delta A
CC   elements, NFKB elements and/or Ikaros-binding oligonucleotides, bind to
CC   delta A elements, NFKB elements and/or Ikaros-binding oligonucleotides,
CC   competitively inhibit binding of naturally occurring Ikaros isoforms to
CC   delta A elements, NFKB elements and/or Ikaros-binding oligonucleotides,
CC   competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
CC   inhibit protein-protein interactions of transcriptional complexes with
CC   naturally occurring Ikaros isoforms, can be used to treat immune system
CC   disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.
CC   Alzheimer's disease. The present sequence represents an Ikaros protein
CC   general formula from the present invention.
XX
SQ   Sequence   470 AA;
```

```
alignment_scores:
            Quality: 2207.50              Length:    468
              Ratio:  5.098                Gaps:      3
 Percent Similarity:  92.521   Percent Identity: 90.385

alignment_block:
 US-08-711-417C-165 x AAW72678   ..

 Align seg 1/1  to: AAW72678  from: 1  to: 470

     160 GCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAATGGGCGTGC 209
         ||||||||||||||||||||||||||||||||||||||||||||||||||
       3 AlaSerAsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAl 19

     210 CTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAATGCTTGATG 259
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      19 aCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspA 36

     260 CCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGCAGCTCGGCT 309
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      36 laSerGlyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAla 52

     310 TTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACTAAAGTGTGA 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      53 LeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAs 69

     360 TATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAAA 409
         |||||||||||||   ||||||||||||||||||||||||||||||||||
      69 pIleCysGlyIle***CysIleGlyProAsnValLeuMetValHisLysA 86

     410 GAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGCGGGGCCTCA 459
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      86 rgSerHisThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSer 102

     460 TTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGGGA 509
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     103 PheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGl 119

     510 GAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCCGGAGGGACG 559
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     119 uLysProPheLysCysHisLeuCysAsnTyrAlaCysArgArgArgAspA 136

     560 CCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATGT 609
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     136 laLeuThrGlyHisLeuArgThrHisSerValGlyLysProHisLysCys 152

     610 GGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGAGGAACATAA 659
         ||||||||||||||||||||||||||||||:::|||||||||||||||||
     153 GlyTyrCysGlyArgSerTyrLysGlnArg***SerLeuGluGluHisLy 169

     660 AGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGGGCACACTGT 709
         ||||||||||||||||||||||||||||||||||||||||||||:::
     169 sGluArgCysHisAsnTyrLeuGluSerMetGlyLeuProGly******* 186

     710 ACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCAGAAGACCTG 759
         ||||||||||||||||||||   |||:::||||||||||||||||||||
     186 **ProValIleLysGluGluThr***His***GluMetAlaGluAspLeu 202

     760 TGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACTAGCAAGTAA 809
         ||||||||||||:::|||||||||||||||||||||||||||||||||||
     203 CysLysIleGly***GluArgSerLeuValLeuAspArgLeuAlaSerAs 219

     810 TGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGACAAGG 859
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     219 nValAlaLysArgLysSerSerMetProGlnLysPheLeuGlyAspLys* 236

     860 GCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAGGAGAACGAA 909
         |||||||||:::|||||||||||||||:::||||||||||||
     236 **LeuSerAsp***ProTyrAspSerAla***TyrGluLysGlu****** 252

     910 ATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTA 959
         ||||||   ||||||||||||||   ||||||||||||||||||||||||
```

```
     253 MetMet***SerHisValMetAsp***AlaIleAsnAsnAlaIleAsnTy 269

     960 CCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCCCGGGCGGTT 1009
         ||||||||||||||||||||||||||||||||||||||||||||||||   |
     269 rLeuGlyAlaGluSerLeuArgProLeuValGlnThrProProGly***S 286

    1010 CCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC...AGGCGC 1056
         ||||||||||||||||||||||||||||||||||||||||||
     286 erGluValValProValIleSerProMetTyrGlnLeuHis********* 302

    1057 TCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAGCGCCGTGGA 1106
         |||   |||:::|||||||||||||||||||||||||||||:::||||||
     303 Ser***Gly***ProArgSerAsnHisSerAlaGlnAsp***AlaVal** 319

    1107 GTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGGAGCGCGAGG 1156
             ||||||||||||||||||||||   |||   ||||||||||||
     319 ****LeuLeuLeuLeuSerLysAlaLys***Val***SerGluArgGluA 336

    1157 CGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAGAGCAACAAC 1206
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     336 laSerProSerAsnSerCysGlnAspSerThrAspThrGluSerAsn*** 352

    1207 GAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACATCGCCCGACG 1256
         ||||||||||||||||||||||||||||||||||||||||||||:::
     353 GluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIle******** 369

    1257 CGCG...CAACGCGTGTCGCTCAAGGAGGAGCACCGCGCCTACGACCTGC 1303
          |||          :::|||||||||||||   |||||||||      |
     369 *Ala***************LeuLysGluGlu***ArgAlaTyr******L 386

    1304 TGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGCACC 1353
         ||||||||||||||||||||||||||||||||   ||||||||||||||||
     386 euArgAlaAlaSerGluAsnSerGlnAspAla***ArgValValSerThr 402

    1354 AGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCGGGTGCTCTT 1403
         ||||||||||||   |||||||||||||||||||||||||||||||||||
     403 SerGlyGluGln***LysValTyrLysCysGluHisCysArgValLeuPh 419

    1404 CCTGGATCACGTCATGTACACCATCCACATG.........GGCTGCCACG 1444
         ||||||||||||||||||||||||||||||||        ||||||||||
     419 eLeuAspHisValMetTyrThrIleHisMet*********GlyCysHisG 436

    1445 GCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCACAGCCAGGAC 1494
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     436 lyPheArgAspProPheGluCysAsnMetCysGlyTyrHisSerGlnAsp 452

    1495 CGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGCTTCCACAT 1544
         ||||||||||||||||||||||||||||||||||||||||||   |||
     453 ArgTyrGluPheSerSerHisIleThrArgGlyGluHisArg***His** 469

    1545 GAGC 1548
          |||
     469 *Ser 470

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW70970

seq_documentation_block:
ID   AAW70970 standard; Protein; 470 AA.
XX
AC   AAW70970;
XX
DT   11-JAN-1999  (first entry)
XX
DE   Ikaros isoform 1 consensus.
XX
KW   Ikaros; mIK-1; hIk-1; transcription factor; mouse; human;
KW   lymphocyte; cell differentiation; T cell; cancer;
KW   immunodeficiency; Alzheimer's disease; therapy; diagnosis.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
```

```
FT                   /note= "variable"
FT   Misc-difference 2
FT                   /note= "variable"
FT   Misc-difference 74
FT                   /label= Gly, Ala, Val, Ile, Leu, Ser, Thr
FT   Misc-difference 89
FT                   /note= "the codon for Thr-89 may be made
FT                          degenerate to provide a stop codon in
FT                          recombinant genes of a degenerate library"
FT   Misc-difference 145
FT                   /note= "the codon for Ser-145 may be made
FT                          degenerate to provide a stop codon in
FT                          recombinant genes of a degenerate library"
FT   Misc-difference 163
FT                   /note= "variable"
FT   Misc-difference 184
FT                   /label= Gly, Ala, Val, Ile, Leu, Ser, Thr
FT                   /note= "residue 184 may also not be present"
FT   Misc-difference 185
FT                   /note= "variable"
FT   Misc-difference 186
FT                   /note= "variable"
FT   Misc-difference 187
FT                   /note= "the codon for Pro-187 may be made
FT                          degenerate to provide a stop codon in
FT                          recombinant genes of a degenerate library"
FT   Misc-difference 194
FT                   /note= "variable"
FT   Misc-difference 196
FT                   /note= "variable"
FT   Misc-difference 207
FT                   /note= "variable"
FT   Misc-difference 232
FT                   /note= "the codon for Gly-232 may be made
FT                          degenerate to provide a stop codon in
FT                          recombinant genes of a degenerate library"
FT   Misc-difference 236
FT                   /note= "variable"
FT   Misc-difference 240
FT                   /note= "variable"
FT   Misc-difference 246
FT                   /note= "variable"
FT   Misc-difference 251
FT                   /note= "variable"
FT   Misc-difference 252
FT                   /note= "variable"
FT   Misc-difference 255
FT                   /note= "variable"
FT   Misc-difference 261
FT                   /note= "variable"
FT   Misc-difference 285
FT                   /note= "variable"
FT   Misc-difference 300
FT                   /note= "variable"
FT   Misc-difference 301
FT                   /note= "variable"
FT   Misc-difference 302
FT                   /note= "variable"
FT   Misc-difference 304
FT                   /note= "variable"
FT   Misc-difference 306
FT                   /note= "variable"
FT   Misc-difference 316
FT                   /note= "variable"
FT   Misc-difference 319
FT                   /note= "variable"
FT   Misc-difference 320
FT                   /note= "variable"
FT   Misc-difference 329
FT                   /note= "variable"
FT   Misc-difference 331
FT                   /note= "variable"
FT   Misc-difference 352
FT                   /note= "variable"
FT   Misc-difference 367
FT                   /note= "variable"
FT   Misc-difference 368
FT                   /note= "variable"
FT   Misc-difference 369
FT                   /note= "variable"
FT   Misc-difference 371
FT                   /note= "variable"
FT   Misc-difference 372
FT                   /note= "variable"
FT   Misc-difference 373
FT                   /note= "variable"
FT   Misc-difference 374
FT                   /note= "variable"
FT   Misc-difference 375
FT                   /note= "variable"
FT   Misc-difference 380
FT                   /note= "variable"
FT   Misc-difference 384
FT                   /note= "variable"
FT   Misc-difference 385
FT                   /note= "variable"
FT   Misc-difference 397
FT                   /note= "variable"
FT   Misc-difference 407
FT                   /note= "variable"
FT   Misc-difference 430
FT                   /note= "variable"
FT   Misc-difference 431
FT                   /note= "variable"
FT   Misc-difference 432
FT                   /note= "variable"
FT   Misc-difference 467
FT                   /note= "variable"
FT   Misc-difference 468
FT                   /note= "variable"
XX
PN   CA2194256-A.
XX
PD   05-MAR-1998.
XX
PF   02-JAN-1997;   97CA-2194256.
XX
PR   05-SEP-1996;   96US-0711417.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1998-378292/33.
XX
PT   New nucleic acid encoding Ikaros protein involved in early
PT   differentiation of lymphocytes - existing in several isoforms, and
PT   related products, used to treat e.g. immune diseases or cancer and
PT   to control cell differentiation
XX
PS   Disclosure; Page 59-60; 158pp; English.
XX
CC   This is an example of a potential Ikaros sequence derived from a
CC   degenerate library of polypeptides based on the amino acid
CC   sequences of human and murine Ikaros isoform 1 proteins hIk-1 (see
CC   AAW70964) and mIk-1 (see AAW70966).  A combinatorial library is
CC   produced using a degenerate library of genes which each include
CC   at least a portion of potential Ikaros sequences.  It can be
CC   generated by combinatorial mutagenesis at the nucleic acid level.
CC   Native Ikaros is active in the early stages of lymphocyte
CC   differentiation.  Different isoforms arise from differential
CC   splicing of Ikaros gene transcripts.  They are expressed primarily
CC   in T cells in the adult and may play a role as a genetic switch
CC   regulating entry into the T cell lineage.  The invention provides
CC   Ikaros nucleic acids (see AAV42805-11 and AAV42840), polypeptides (see
CC   AAW70963-71), vectors and host cells.  These can be used to treat T
```

```
CC   and B cell diseases (e.g. immune deficiencies caused by drugs,
CC   radiation or cancer), to control expression of heterologous genes
CC   placed under control of an Ikaros-responsive element, to treat
CC   nervous system diseases (e.g. Alzheimer's disease) and to modulate
CC   cell division, amplification or differentiation, especially
CC   in haematopoietic cells.  Some Ikaros isoforms are antagonistic of
CC   others and may be used to inhibit interaction with DNA sequences.
XX
SQ   Sequence   470 AA;


alignment_scores:
           Quality: 2204.50            Length:    468
             Ratio:  5.091              Gaps:      3
 Percent Similarity:  92.521   Percent Identity: 90.171

alignment_block:
 US-08-711-417C-165 x AAW70970   ..

 Align seg 1/1  to: AAW70970  from: 1  to: 470

     160 GCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAATGGGCGTGC 209
         |||||||||||||||||||||||||||||||||||||||||||||||||||
       3 AlaSerAsnValLysValGluThrGlnSerAspGluGluAsnGlyArgAl 19

     210 CTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAATGCTTGATG 259
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      19 aCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgMetLeuAspA 36

     260 CCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGCAGCTCGGCT 309
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      36 laSerGlyGluLysMetAsnGlySerHisArgAspGlnGlySerSerAla 52

     310 TTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACTAAAGTGTGA 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      53 LeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLeuLysCysAs 69

     360 TATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAAA 409
         |||||||||||||   |||||||||||||||||||||||||||||||||
      69 pIleCysGlyIle***CysIleGlyProAsnValLeuMetValHisLysA 86

     410 GAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGCGGGGCCTCA 459
         |||||||||||||||||||||||||||||||||||||||||||||||||||
      86 rgSerHisThrGlyGluArgProPheGlnCysAsnGlnCysGlyAlaSer 102

     460 TTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCATTCCGGGGA 509
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     103 PheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHisSerGlyGl 119

     510 GAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCCGGAGGGACG 559
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     119 uLysProPheLysCysHisLeuCysAsnTyrAlaCysArgArgArgAspA 136

     560 CCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCTCACAAATGT 609
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     136 laLeuThrGlyHisLeuArgThrHisSerValGlyLysProHisLysCys 152

     610 GGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGAGGAACATAA 659
         |||||||||||||||||||||||||||||:::|||||||||||||||||
     153 GlyTyrCysGlyArgSerTyrLysGlnArg***SerLeuGluGluHisLy 169

     660 AGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGGGCACACTGT 709
         |:::|||||||||||||||||||||||||||||||||||||||||:::
     169 sGlnArgCysHisAsnTyrLeuGluSerMetGlyLeuProGly******* 186

     710 ACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCAGAAGACCTG 759
           ||||||||||||||||||   |||:::|||||||||||||||||||
     186 **ProValIleLysGluGluThr***His***GluMetAlaGluAspLeu 202

     760 TGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACTAGCAAGTAA 809
         ||||||||||||:::|||||||||||||||||||||||||||||||||||
     203 CysLysIleGly***GluArgSerLeuValLeuAspArgLeuAlaSerAs 219

     810 TGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTGGGGACAAGG 859
         |||||||||||||||||||||||||||||||||||||||||||||||||||
     219 nValAlaLysArgLysSerSerMetProGlnLysPheLeuGlyAspLys* 236

     860 GCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAGGAGAACGAA 909
         |||||||||:::|||||||||||||||||:::|||||||||||||
     236 **LeuSerAsp***ProTyrAspSerAla***TyrGluLysGlu****** 252

     910 ATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGCCATCAACTA 959
         ||||||   ||||||||||||||||   ||||||||||||||||||||||
     253 MetMet***SerHisValMetAsp***AlaIleAsnAsnAlaIleAsnTy 269

     960 CCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCCCGGGCGGTT 1009
         |||||||||||||||||||||||||||||||||||||||||||||||||   |
     269 rLeuGlyAlaGluSerLeuArgProLeuValGlnThrProProGly***S 286

    1010 CCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC...AGGCGC 1056
         |||||||||||||||||||||||||||||||||||||||||
     286 erGluValValProValIleSerProMetTyrGlnLeuHis********* 302

    1057 TCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAGCGCCGTGGA 1106
         |||   |||:::|||||||||||||||||||||||||||:::||||||
     303 Ser***Gly***ProArgSerAsnHisSerAlaGlnAsp***AlaVal** 319

    1107 GTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGGAGCGCGAGG 1156
              ||||||||||||||||||||||||   |||.  ||||||||||||
     319 ****LeuLeuLeuLeuSerLysAlaLys***Val***SerGluArgGluA 336

    1157 CGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAGAGCAACAAC 1206
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     336 laSerProSerAsnSerCysGlnAspSerThrAspThrGluSerAsn*** 352

    1207 GAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACATCGCCCGACG 1256
         ||||||||||||||||||||||||||||||||||||||||||:::
     353 GluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIle******** 369

    1257 CGCG...CAACGCGTGTCGCTCAAGGAGGAGCACCGCGCCTACGACCTGC 1303
          |||            :::|||||||||||   |||||||||      |
     369 *Ala***************LeuLysGluGlu***ArgAlaTyr*******L 386

    1304 TGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGCACC 1353
         |||||||||||||||||||||||||||||||   ||||||||||||||||
     386 euArgAlaAlaSerGluAsnSerGlnAspAla***ArgValValSerThr 402

    1354 AGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCGGGTGCTCTT 1403
         ||||||||||||   ||||||||||||||||||||||||||||||||||
     403 SerGlyGluGln***LysValTyrLysCysGluHisCysArgValLeuPh 419

    1404 CCTGGATCACGTCATGTACACCATCCACATG.........GGCTGCCACG 1444
         |||||||||||||||||||||||||||||||         ||||||||||
     419 eLeuAspHisValMetTyrThrIleHisMet*********GlyCysHisG 436

    1445 GCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCACAGCCAGGAC 1494
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     436 lyPheArgAspProPheGluCysAsnMetCysGlyTyrHisSerGlnAsp 452

    1495 CGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGCTTCCACAT 1544
         ||||||||||||||||||||||||||||||||||||||||||   |||
     453 ArgTyrGluPheSerSerHisIleThrArgGlyGluHisArg***His** 469

    1545 GAGC 1548
          |||
     469 *Ser 470

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR92016

seq_documentation_block:
ID   AAR92016 standard; Protein; 432 AA.
XX
AC   AAR92016;
XX
DT   08-MAY-1996  (first entry)
XX
```

```
DE   Murine Ikaros protein mIk-3.
XX
KW   Ikaros; transgene; transgenic animal; transgenic mouse; lymphocyte;
KW   immunocomprised; immune system disorder; nervous system disorder;
KW   animal model; mIk-3.
XX
OS   Mus musculus.
XX
PN   WO9604372-A1.
XX
PD   15-FEB-1996.
XX
PF   28-JUL-1995;   95WO-US09345.
XX
PR   29-JUL-1994;   94US-0283300.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1996-129389/13.
DR   N-PSDB; AAT16061.
XX
PT   Transgenic rodent having Ikaros trans-gene (pref. mutated)  - is
PT   severely immuno-compromised and can be used as model to determine
PT   effects of treatment for immune and nervous system disorders
XX
PS   Disclosure; Page 63-64; 102pp; English.
XX
CC   The sequence of 48 kDa mouse Ikaros protein mIK-3 (AAR92016) was
CC   deduced from mouse Ikaros cDNA (AAT16061) isolated from a mature
CC   T-cell line E15 library.  Ikaros protein is a master regulator of
CC   hematopoietic differentiation and a major determinant in lymphocyte
CC   differentiation.  Other isoforms of Ikaros (see AAR92014 and
CC   AAR92017-19) arise from differential splicing of Ikaros gene
CC   transcripts.  Transgenic animals, esp. mice, having a mutated Ikaros
CC   transgene, esp. a mutation that alters the DNA binding domain of the
CC   Ikaros protein, are used as models to determine the effects of
CC   treatments for immune or nervous system disorders.
XX
SQ   Sequence   432 AA;


alignment_scores:
            Quality: 1963.00              Length:    521
              Ratio:  4.776                Gaps:      6
 Percent Similarity:  78.887   Percent Identity: 74.280

alignment_block:
 US-08-711-417C-165 x AAR92016   ..

 Align seg 1/1  to: AAR92016  from: 1  to: 432

       1 ATGGATGCTGACGAGGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG 50
         ||||||:::|||||||||||||||||||      |||||||||||||||
       1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17

      51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
         ||||||||||||||||||||||||||||||||||||||||||||||:::|
      17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34

     101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAGCTCCAAGAGTGAC 150
         |||||||||||||||||||||||:::||||||:::||||||||||||
      34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50

     151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAA 200
         |||   :::||||||||||||||||||||||||||||||||||||||||
      51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67

     201 TGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAA 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84

     251 TGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100

     301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117

     351 AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400
         ||||||||||||||||||||:::|||||||||||||||||||||||||||
     117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134

     401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnSer 150

     451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCA 500
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     151 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 167

     501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCC 550
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     167 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 184

     551 GGAGGGACGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
         ||||||||||||||||||||||||||||||||||||
     184 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSer..... 196

     601 CACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650

     196 ..................................................... 196

     651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700

     196 ..................................................... 196

     701 GCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCA 750

     196 ..................................................... 196

     751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800

     196 ..................................................... 196

     801 AGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
                                                           |
     197 .....................................................G 197

     851 GGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAG 900
         ||||||||   ||||||||   ||||||||||||||:::|||||||||
     197 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 213

     901 GAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
         |||   :::||||||   |||||||||||||||||||||||||||||||
     214 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl 229

     951 CATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCC 1000
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     229 aIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProP 246

    1001 CGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC 1050
         |||||:::|||||||||||||||||||||||   ||||||||||||||
     246 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 262

    1051 AGG...CGCTCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAG 1097
         :::      |||:::|||   ||||||||||||||||||||||||||
     263 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp.. 278

    1098 CGCCGTGGAGTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGG 1147
          ||||||:::   ||||||||||||||||||||||   |||   ||||
     279 .AlaValAspAsnLeuLeuLeuLeuSerLysAlaLysSerValSerSerG 295

    1148 AGCGCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG 1197
```

```
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 295 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 311

1198 AGCAACAACGAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACAT 1247
     ||||||   ||||||||||||||||||||||||||||||||||||||||
 312 SerAsnAlaGluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIl 328

1248 CGCCCGACGCGCGCAACGC...GTGTCGCTCAAGGAGGAGCACCGCGCCT 1294
     |      :::|||:::::   ::::::|||||||||||||:::|||||||
 328 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluGlnArgAlaT 345

1295 ACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTG 1344
     ||:::::|||||||||||||||||||||||||||||||:::|||||||
 345 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 361

1345 GTCAGCACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCG 1394
     ||||||||||||||||||:::|||||||||||||||||||||||||||
 362 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 378

1395 GGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATG.........G 1435
     |||||||||||||||||||||||||||||||||||||||         |
 378 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 395

1436 GCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCAC 1485
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 395 lyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 411

1486 AGCCAGGACCGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCG 1535
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 412 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 428

1536 CTTCCACATGAGC 1548
     |:::|||:::|||
 428 gTyrHisLeuSer 432
```

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:AAW72673

seq_documentation_block:

```
ID   AAW72673 standard; Protein; 432 AA.
XX
AC   AAW72673;
XX
DT   14-JAN-1999  (first entry)
XX
DE   Mouse Ikaros mIk-3.
XX
KW   CD3-delta gene; Ikaros gene; T cell; progenitor stem cell; leukaemia;
KW   differentiation marker; immune system; corpus striatum; AIDS;
KW   Alzheimer's disease.
XX
OS   Mus sp.
XX
PN   US5824770-A.
XX
PD   20-OCT-1998.
XX
PF   05-JUN-1995;   95US-0465590.
XX
PR   02-MAY-1994;   94US-0238212.
PR   14-SEP-1992;   92US-0946233.
PR   14-SEP-1993;   93US-0121438.
PR   05-JUN-1995;   95US-0465590.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Georgopoulos K;
XX
DR   WPI; 1998-582621/49.
DR   N-PSDB; AAV66970.
XX
PT   Ikaros poly:peptide(s) - useful for treating disorders of immune
PT   system or corpus striatum
XX
PS   Claim 1; Column 57-62; 111pp; English.
XX
CC   The present invention describes a purified peptide having at least one
CC   of the following properties: (a) it stimulates transcription of a DNA
CC   sequence under the control of a delta A element, an NFKB element or an
CC   Ikaros binding oligonucleotide consensus sequence; (b) it binds to any of
CC   a delta A element, an NFKB element or an Ikaros binding oligonucleotide
CC   consensus sequence; (c) it competitively inhibits the binding of a
CC   naturally occurring Ikaros isoform to any of a delta A element, an NFKB
CC   element or an Ikaros binding oligonucleotide consensus sequence; (d) it
CC   competitively inhibits Ikaros binding to Ikaros responsive elements; or
CC   (e) it inhibits protein-protein interactions of transcriptional complexes
CC   formed with naturally occurring Ikaros isoforms. The proteins, provided
CC   that they stimulate gene transcription under the control of delta A
CC   elements, NFKB elements and/or Ikaros-binding oligonucleotides, bind to
CC   delta A elements, NFKB elements and/or Ikaros-binding oligonucleotides,
CC   competitively inhibit binding of naturally occurring Ikaros isoforms to
CC   delta A elements, NFKB elements and/or Ikaros-binding oligonucleotides,
CC   competitively inhibit Ikaros binding to Ikaros-responsive elements and/or
CC   inhibit protein-protein interactions of transcriptional complexes with
CC   naturally occurring Ikaros isoforms, can be used to treat immune system
CC   disorders, e.g. leukaemia or AIDS, or corpus striatum disorders, e.g.
CC   Alzheimer's disease. The present sequence represents a specifically
CC   claimed mouse Ikaros protein.
XX
SQ   Sequence   432 AA;
```

```
alignment_scores:
           Quality: 1963.00              Length:    521
             Ratio:  4.776                Gaps:      6
 Percent Similarity:  78.887   Percent Identity: 74.280

alignment_block:
 US-08-711-417C-165 x AAW72673   ..

 Align seg 1/1  to: AAW72673  from: 1  to: 432

       1 ATGGATGCTGACGAGGGTCAAGACATGTCTTTCTCATCAGGGAAGGAAAG 50
         ||||||:::|||||||||||||||||||||      ||||||||||||||
       1 MetAspValAspGluGlyGlnAspMetSerGlnValSerGlyLysGluSe 17

      51 CCCCCCTGTAAGCGATACTCCAGATGAGGGCGATGAGCCCATGCCGATCC 100
         ||||||||||||||||||||||||||||||||||||||||||||||:::|
      17 rProProValSerAspThrProAspGluGlyAspGluProMetProValP 34

     101 CCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAGCTCCAAGAGTGAC 150
         |||||||||||||||||||||||||:::|||||||:::|||||||||||||
      34 roGluAspLeuSerThrThrSerGlyAlaGlnGlnAsnSerLysSerAsp 50

     151 AGAGTCGTGGCCAGTAATGTTAAAGTAGAGACTCAGAGTGATGAAGAGAA 200
         |||   :::|||||||||||||||||||||||||||||||||||||||||
       51 ArgGlyMetAlaSerAsnValLysValGluThrGlnSerAspGluGluAs 67

     201 TGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTACGAA 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      67 nGlyArgAlaCysGluMetAsnGlyGluGluCysAlaGluAspLeuArgM 84

     251 TGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      84 etLeuAspAlaSerGlyGluLysMetAsnGlySerHisArgAspGlnGly 100

     301 AGCTCGGCTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACT 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     101 SerSerAlaLeuSerGlyValGlyGlyIleArgLeuProAsnGlyLysLe 117

     351 AAAGTGTGATATCTGTGGGATCATTTGCATCGGGCCCAATGTGCTCATGG 400
         ||||||||||||||||||||||:::|||||||||||||||||||||||||
     117 uLysCysAspIleCysGlyIleValCysIleGlyProAsnValLeuMetV 134

     401 TTCACAAAAGAAGCCACACTGGAGAACGGCCCTTCCAGTGCAATCAGTGC 450
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     134 alHisLysArgSerHisThrGlyGluArgProPheGlnCysAsnGlnSer 150
```

```
 451 GGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCA 500
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 151 GlyAlaSerPheThrGlnLysGlyAsnLeuLeuArgHisIleLysLeuHi 167

 501 TTCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCC 550
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 167 sSerGlyGluLysProPheLysCysHisLeuCysAsnTyrAlaCysArgA 184

 551 GGAGGGACGCCCTCACTGGCCACCTGAGGACGCACTCCGTTGGTAAACCT 600
     ||||||||||||||||||||||||||||||||||
 184 rgArgAspAlaLeuThrGlyHisLeuArgThrHisSer..... 196

 601 CACAAATGTGGATATTGTGGCCGAAGCTATAAACAGCGAACGTCTTTAGA 650

 196 ..................................................  196

 651 GGAACATAAAGAGCGCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGG 700

 196 ..................................................  196

 701 GCACACTGTACCCAGTCATTAAAGAAGAAACTAAGCACAGTGAAATGGCA 750

 196 ..................................................  196

 751 GAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTCGTGCTGGACAGACT 800

 196 ..................................................  196

 801 AGCAAGTAATGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTTCTTG 850
                                                        |
 197 .....................................................G 197

 851 GGGACAAGGGCCTGTCCGACACGCCCTACGACAGTGCCACGTACGAGAAG 900
     ||||||||   |||||||||   ||||||||||||||:::|||||||||
 197 lyAspLysCysLeuSerAspMetProTyrAspSerAlaAsnTyrGluLys 213

 901 GAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACGC 950
     |||   :::||||||   ||||||||||||||||||||||||||||||||
 214 Glu...AspMetMetThrSerHisValMetAspGlnAlaIleAsnAsnAl 229

 951 CATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCC 1000
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 229 aIleAsnTyrLeuGlyAlaGluSerLeuArgProLeuValGlnThrProP 246

1001 CGGGCGGTTCCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCAC 1050
     ||||||:::|||||||||||||||||||||||   ||||||||||||||
 246 roGlySerSerGluValValProValIleSerSerMetTyrGlnLeuHis 262

1051 AGG...CGCTCGGAGGGCACCCCGCGCTCCAACCACTCGGCCCAGGACAG 1097
     :::      |||:::|||   ||||||||||||||||||||||||||
 263 LysProProSerAspGlyProProArgSerAsnHisSerAlaGlnAsp.. 278

1098 CGCCGTGGAGTACCTGCTGCTGCTCTCCAAGGCCAAGTTGGTGCCCTCGG 1147
      ||||||:::   |||||||||||||||||||||||   |||   ||||
 279 .AlaValAspAsnLeuLeuLeuLeuSerLysAlaLysSerValSerSerG 295

1148 AGCGCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCACGGACACCGAG 1197
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 295 luArgGluAlaSerProSerAsnSerCysGlnAspSerThrAspThrGlu 311

1198 AGCAACAACGAGGAGCAGCGCAGCGGTCTTATCTACCTGACCAACCACAT 1247
     ||||||   |||||||||||||||||||||||||||||||||||||||||
 312 SerAsnAlaGluGluGlnArgSerGlyLeuIleTyrLeuThrAsnHisIl 328

1248 CGCCCGACGCGCGCAACGC...GTGTCGCTCAAGGAGGAGCACCGCGCCT 1294
     |      :::|||::::::   ::::::||||||||||||:::|||||||
 328 eAsnProHisAlaArgAsnGlyLeuAlaLeuLysGluGluGlnArgAlaT 345

1295 ACGACCTGCTGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTG 1344
     ||::::::|||||||||||||||||||||||||||||||:::||||||
 345 yrGluValLeuArgAlaAlaSerGluAsnSerGlnAspAlaPheArgVal 361

1345 GTCAGCACCAGCGGGGAGCAGATGAAGGTGTACAAGTGCGAACACTGCCG 1394
     ||||||||||||||||||||:::||||||||||||||||||||||||||
 362 ValSerThrSerGlyGluGlnLeuLysValTyrLysCysGluHisCysAr 378

1395 GGTGCTCTTCCTGGATCACGTCATGTACACCATCCACATG.........G 1435
     ||||||||||||||||||||||||||||||||||||||         |
 378 gValLeuPheLeuAspHisValMetTyrThrIleHisMetGlyCysHisG 395

1436 GCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACCAC 1485
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 395 lyCysHisGlyPheArgAspProPheGluCysAsnMetCysGlyTyrHis 411

1486 AGCCAGGACCGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCG 1535
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 412 SerGlnAspArgTyrGluPheSerSerHisIleThrArgGlyGluHisAr 428

1536 CTTCCACATGAGC 1548
     |:::|||:::|||
 428 gTyrHisLeuSer 432
```